

SEARCH REQUEST FORM

2-631

Requestor's Name: _____ Serial Number: _____

Date: _____ Phone: _____ Art Unit: 9E11

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 2/1/79

Searcher: Sheppard

Terminal time: _____

Elapsed time: _____

CPU time: _____

Total time: _____

Number of Searches: _____

Number of Databases: _____

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

_____ Other

W P E R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 14:40:49 1999; MasPar time 2624.06 Seconds
Abular output not generated. 1532.792 Million cell updates/sec

Title: >US-08-934-254-26
Description: (1-1702) from US08934254.seq
Perfect Score: 1702
N.A. Sequence: 1 CCCCAAAATTTTCATTGTT.....TTTTGGTAAAAAAAAA 1702
Comp: GGGGTTTTTAAAGTAACAA.....AAAAACCAATTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb157

1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_vi

Database: genbank110
16:gb_bal 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pr1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_vi

Statistics: Mean 11.427; Variance 5.921; scale 1.930

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
SUMMARIES							
c 1	399	23.4	110149	25	ATAC005397	Arabidopsis thaliana c	4.77e-288
2	393	23.1	1705	25	ATAJ4161	Arabidopsis thaliana m	3.71e-283
3	385	22.6	1610	25	BNJ4160	Brassica napus mRNA fo	1.22e-276
4	340	20.0	1591	25	HACTYBERN	H.annuus mRNA for extr	4.63e-240
5	172	10.1	1685	22	AR020904	Sequence 4 from patent	1.78e-105
6	172	10.1	1685	22	I38430	Sequence 4 from patent	1.78e-105
7	172	10.1	1687	25	BOU79010	Borago officinalis del	1.78e-105
8	91	5.3	7218	22	166494	Sequence 14 from paten	2.20e-43
9	46	2.7	965	22	AR024229	Sequence 22 from paten	7.60e-12
10	43	2.5	281	24	RYENRRA	Secale cereale nitrate	6.13e-10
c 11	42	2.5	10772	19	AF012089	Drosophila melanogaste	2.59e-09
c 12	40	2.4	965	22	AR024229	Sequence 22 from paten	4.48e-08
13	40	2.4	10772	19	AF012089	Drosophila melanogaste	4.48e-08

c 14	37	2.2	7218	22	I66494	Sequence 14 from paten	2.92e-06
c 15	38	2.2	74371	27	AC005369	Homo sapiens chromosom	7.36e-07
c 16	37	2.2	74371	27	AC005369	Homo sapiens chromosom	2.92e-06
c 17	36	2.1	215	22	I28278	Sequence 5 from patent	1.14e-05
c 18	35	2.1	285	24	BLYNRB	Hordeum chilense nitra	4.41e-05
c 19	35	2.1	285	24	BLYNRC	Hordeum lechleri nitra	4.41e-05
c 20	35	2.1	285	24	BLYNRD	Hordeum pusillum nitra	4.41e-05
c 21	36	2.1	1801	24	MZENADNR	Zea mays nitrate reduc	1.14e-05
c 22	35	2.1	3076	24	CUCNITRA	C.maxima nitrate reduc	4.41e-05
c 23	35	2.1	3538	24	HVNAR7	Hordeum vulgare nar7 g	4.41e-05
c 24	36	2.1	4985	24	ZMU20450	Zea mays nitrate reduc	1.14e-05
c 25	35	2.1	7602	24	VCNITA	V.cartieri gene for nit	4.41e-05
c 26	34	2.0	215	22	I28278	Sequence 5 from patent	1.67e-04
c 27	34	2.0	2000	24	OSNIA134	Rice nial gene for nit	1.67e-04
c 28	33	1.9	282	24	BLYNRA	Hordeum bogdanii nitra	6.22e-04
c 29	33	1.9	285	24	BLYNRF	Hordeum stenostachys n	6.22e-04
c 30	32	1.9	2126	24	MZENAR	Maize NADH:nitrate red	2.27e-03
c 31	33	1.9	5499	24	LJNIA	L.japonicus NIA gene.	6.22e-04
c 32	31	1.8	273	24	ASTNRA	Avena strigosa nitrate	8.15e-03
c 33	31	1.8	540	22	E13060	cDNA encoding cytochro	8.15e-03
c 34	31	1.8	565	22	E04076	gDNA encoding envelope	8.15e-03
c 35	30	1.8	3811	26	HUMIMUCA	Human mucin 2 (MUC2) m	2.86e-02
c 36	30	1.8	15720	27	HUMMUC2X	Homo sapiens intestina	2.86e-02
c 37	30	1.8	112309	27	AC003025	Human Chromosome 11p12	2.86e-02
c 38	30	1.8	170743	18	AC004228	*** SEQUENCING IN PROG	2.86e-02
c 39	30	1.8	177014	27	AC004063	Homo sapiens chromosom	2.86e-02
c 40	29	1.7	954	24	CVNIREDB	C.vulgaris mRNA for ni	9.84e-02
c 41	29	1.7	38962	16	SC2F1	Streptomyces coelicolo	9.84e-02
c 42	29	1.7	75342	19	CEY47H9C	Caenorhabditis elegans	9.84e-02
c 43	29	1.7	110000	18	CEY53C10	Caenorhabditis elegans	9.84e-02
c 44	29	1.7	181098	27	AC004216	Homo sapiens 12q24 PAC	9.84e-02
c 45	29	1.7	231298	18	AC005921	*** SEQUENCING IN PROG	9.84e-02

ALIGNMENTS

RESULT	1	ATAC005397	110149 bp	DNA	PLN	06-OCT-1998
LOCUS		Arabidopsis thaliana chromosome II BAC T3F17	genomic sequence,			
DEFINITION		complete sequence.				
ACCESSION		AC005397				
NID		93702315				
KEYWORDS		HTG.				
SOURCE		thale cress.				
ORGANISM		Arabidopsis thaliana				
REFERENCE		1 (bases 1 to 110149)				
AUTHORS		Rounsley, S.D., Lin, X., Kaul, S., Shea, T.P., Fujii, C.Y., Mason, T.M., Shen, M., Ronning, C.M., Fraser, C.M., Somerville, C.R. and Venter, J.C.				
TITLE		Arabidopsis thaliana chromosome II BAC T3F17 genomic sequence				
JOURNAL		Unpublished				
REFERENCE		2 (bases 1 to 110149)				
AUTHORS		Rounsley, S.D. and Lin, X.				
TITLE		Direct Submission				
JOURNAL		Submitted (07-AUG-1998)				
REFERENCE		3 (bases 1 to 110149)				
AUTHORS		Rounsley, S.D.				
TITLE		Direct Submission				
JOURNAL		Submitted (06-OCT-1998)				
COMMENT		Medical Center Dr., Rockville, MD 20850, USA				
		On Oct 6, 1998 this sequence version replaced gi:3603083.				
		Address all correspondence to:				
		Steve Rounsley				
		The Institute for Genomic Research				
		9712 Medical Center Dr.,				
		Rockville, MD 20850,				
		USA				
		e-mail: rounsley@tigr.org				
		BAC clone T3F17 is from Arabidopsis chromosome II and is contained				

in the YAC clone C1C02E07.

The orientation of the sequence is from SP6 to T7 end of the BAC clone.

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (available by anonymous ftp from arthur.epm.ornl.gov), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://genomic.stanford.edu/~chris/GENSCANW.html>), and NetPlantGene (<http://www.cds.dcu.ie/netgene/cbsnetpgene.html>), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (<http://www.tigr.org/tdb/at/at.html>).

Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted as tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRAIL are annotated as misc features.

FEATURES

Location/Qualifiers

```

1..110149
  /organism="Arabidopsis thaliana"
  /cultivar="Columbia"
  /db_xref="taxon:3702"
  /chromosome="II"
  /map="C1C02E07"
  /clone="T3F17"
  /gene="T3F17"
  join(<1..21,180..497,585..644,805..>996)
  /gene="T3F17.1"
  <1..>996
  /gene="T3F17.1"
  join(<1..21,180..497,585..644,805..996)
  /gene="T3F17.1"
  /codon_start=1
  /product="putative CopI protein, 5' partial"
  /db_xref="PID:g3702316"
  /translation="KLMSINERSLGTWSPANVCQVQFSSVSNHLLAFGSADYKYVC
  IDLRVYKFPWCLAGHAYSVYKFDSETIVSASTDNLKWNKTNSSGLSPGAC
  SLTYKGNQNRNVLGSLVDGIACSETIVSYIKSLPMPMTYKFGSDVPISGNE
  YFDDNGQFVSSVCKRKNMLVAANSTGNMKLLKIV"
  1464..1565
  /note="exon predicted by xgrail, quality excellent"
  /complement(1490..1522)
  /note="exon predicted by xgrail, quality
  marginal_shadowexon"
  /complement(1857..1921)
  /note="exon predicted by xgrail, quality
  excellent_shadowexon"
  /complement(2157..2251)
  /note="exon predicted by xgrail, quality
  marginal_shadowexon"
  2611..2664
  /note="exon predicted by xgrail, quality marginal"
  2819..2835
  /note="exon predicted by xgrail, quality marginal"
  /complement(3482..3551)
  /note="exon predicted by xgrail, quality good"
  /complement(3773..3849)
  /rpt_family="AT_rich"
  4087..4166
  /rpt_family="POLY_A"
  4106..4166
  /rpt_family="(GAAAA)n"
  /complement(4245..4279)
  /note="exon predicted by xgrail, quality marginal"
  join(4337..4574,4732..5060)
  /gene="T3F17.2"
  4337..5060

```

```

  /gene="T3F17.2"
  join(4460..4574,4732..4838)
  /gene="T3F17.2"
  /note="unknown protein"
  /codon_start=1
  /db_xref="PID:g3702317"
  /translation="MASRNSVTGFALESFVAVILLSAGQSLAPAPAPSDGTSIDQ
  GRAILMVVALVTYLIHPLDAGSSYFF"
  /complement(4798..4921)
  /rpt_family="(GAA)n"
  6186..6362,6458..6628,6761..6811,7054..7176,7264..7341,
  7433..>7570)
  /gene="T3F17.3"
  /complement(<5548..>7570)
  /gene="T3F17.3"
  /complement(join(5548..5677,5751..5809,5945..6094,
  6186..6362,6458..6628,6761..6811,7054..7176,7264..7341,
  7433..>7570))
  /gene="T3F17.3"
  /codon_start=1
  /product="putative mitochondrial carrier protein"
  /db_xref="PID:g3702352"
  /translation="MVVEAQLDPKSKIPNENLKFGERALSGAGAAFIASVINPLDY
  VTRLOQAAGWPYQSGRCGCEDTNLSLVDHRSAPGCRITGSAVCSDNQYKG
  TLDVFYKIIIRSFSTGTLCKRGFLGFGVGIIYPCYDYFRNIMEEFTKSPSLTVTV
  PLVAGTIAESLACISCPVELARTMQAEGTQNRVKLPGWKTLVDVVPVKGSSNG
  YRLMTGLGAQLARDVPFAICWISILEPTRRSIQSANGEPRAGSIIGANFAAGFVAG
  AVAAATCPLDVAKTRROJEKNTDRAMTMTTQTLAEIWRDGGMRGFMFGAGARVGRA
  GPSVAIVSFYEVVKYGLNHFHQ"
  /complement(8132..8179)
  /note="exon predicted by xgrail, quality marginal"
  8202..8275
  /note="exon predicted by xgrail, quality good_shadowexon"
  10697..10738
  /rpt_family="AT_rich"
  /complement(10920..11049)
  /note="exon predicted by xgrail, quality good"
  /complement(<11070..>11954)
  /gene="T3F17.4"
  /complement(<11070..>11954)
  /gene="T3F17.4"
  /complement(11070..11954)
  /gene="T3F17.4"
  /codon_start=1
  /product="putative Ap2 domain protein"
  /db_xref="PID:g3702318"
  /translation="MKSRRKSKYTVHRKITSTPFGFPKIVKIIVTDPCATSSSDE
  ENDKSVAPRVKRYVDEIRFCDEDDPKPAKAKKSPAAAAENGDLVKSVMYKGV
  RQPMGKFAAEIRDPSSRTRLWLTGATAEAAIGYDRAAIRKGNHQAQNTFTPPPS
  PTEVLPTPIDLTETVSGDSASQISLCSPTSVLRFSHNDETYRTPTTEQNPFF
  FLPLDFRSGDYFWDSEITPDPLDFEHQSLLPNINNNVTCKDNLDSFFPLGVIG
  DFSSWDVDEFFQDHLDDK"
  11973..12064
  /rpt_family="(GAAA)n"
  /complement(12493..12534)
  /note="exon predicted by xgrail, quality marginal"
  13238..13306
  /note="exon predicted by xgrail, quality
  marginal_shadowexon"
  /complement(13244..13311)
  /note="exon predicted by xgrail, quality marginal"
  14175..14247
  /rpt_family="(TAAA)n"
  /complement(14264..14323)
  /note="exon predicted by xgrail, quality good"
  /complement(<14412..>15176)
  /gene="T3F17.5"
  /complement(<14412..>15176)
  /note="predicted by genefinder"
  /gene="T3F17.5"
  /complement(14412..15176)
  /gene="T3F17.5"

```


[illegible]

5' UTR
gene

RESULT	3
LOCUS	BNAJ4160 1610 bp mRNA PLN 31-OCT-1998
DEFINITION	Brassica napus mRNA for delta-8 sphingolipid desaturase.

CDS	/gene="sldi" 51..1400 /gene="sldi" /codon_start=1 /evidence=experimental /product="delta-8 sphingolipid desaturase" /db_xref="pid:el337047" /db_xref="pid:g3819708"	
3'UTR	398 a	369 c 375 g 468 t
ORIGIN	Query Match 22.6%; Score 385; DB 25; Length 1610; Best Local Similarity 56.7%; Pred. No. 1.22e-276; Matches 828; Conservative 0; Mismatches 407; Indels 6; Gaps 4;	
Db	102	CACAACCAACCGGAGATTATGATCTCAATCCAAAGGCAAGTCTACGACGTCTCCAC 161
QY	96	CACAACAGTCGGGGATCTCTGGATCTCCATCCAGGCAAGTCTACGACGTCTCCG 155
Db	162	TGGGTCAATCCATCCCGGAGCGAAGCAGCATCTAAACCTGCCGGTCAAGACGTC 221
QY	156	TGGGGGGGAGACACCCCGCGGCGGAGGTCCCGCTCCTCAGTGGCCGGCAGGACGTC 215
Db	222	ACCGACGGCTTCATCGCTTACCATCCCGGAACCGCATCGCGCACCTCGAAACCTCCAC 281
QY	216	ACCGACGGCTTCATCGCTTACCACCGGCGGCGGCGGTGCGGCACTGATCGCTCTTC 275
Db	282	AACGGGTAC--C-ACGTGAAGACCAACACGCTGTCGCGGTGCTGCTGATCAGTACCGTGT 338
QY	276	ACCGGTCTACTACTCTCAAGGACTTCGAAGTGTGCGGCAATCTCAAGGACTACCGGAGG 335
Db	339	TTAGCGGGGAGTTTCCAAACCGGACTCTTCGATATAAAAGGTCACGACTCTTTAC 398
QY	336	CTTTTGAAGAGATGTCCGGTCCGGGTCTTCGAGAAGAGGGCCACCATCATGTGG 395
Db	399	AGCTCACGTGCTGCTGCCATGCTCGCGGCGGTGTATAGCTGTCTGTCATCAGC 458
QY	396	AGCTTCGTCGGGTGCGGTGATGTCGCGGCAATCTCTACGGCTGCTGCGTGGAG 455
Db	459	AGCATATGGCCCACTTAATATCCGCCGCTCTGCTGGGCTTCTCTGGATACAGAGCGCT 518
QY	456	TCCGTGGAGTTACATGCTCTCGCGCGCACTGCTGGGCTTGTGTGGATCCAAAGCGCG 515
Db	519	TAGTGGGACATGACTCTGGTATTACACGTGACGTCAAGACCGGTGTATTAACATC 578
QY	516	TATGTGGGCGATGACTCCGGCCATTACCAAGGTGATGCCAACCCGCTGGATACAGAAATC 575
Db	579	GTCACGCTTCTGCTGGTAAGTCACTACCGGATATCGATCGCGGTGGAAATCGAGC 638
QY	576	AGCAACTCATAGCAGGCAACATCTTAACCGGAATCAGCATCGGTGGGAGTGGAGC 635
Db	639	CATAAGCTTCACCATATCTCTTGTATAGTCTTGACACGATCCTGATCTCCAGCACATC 698
QY	636	CACAGCCCAACCACTCGCTTGCAACGCTCGACTAGACCCCGACCTCCAGCACATC 695
Db	699	CTGTCTTAGCGGTCTCAACAGTTCTTTAAGTCGATGACGTCAAGTCTTCTATGGAGG 758
QY	696	CCCGTATTCGCGTCTCCACCGGACTCTTCAACTCCATCACTCGGTCTCTATGCGCGA 755
Db	759	AGTTGACGTTTCGATCCACTAGCTCGATCTTGATCAGTACCAACACTGGTCTTTAT 818
QY	756	GTCTGAAATTCGAGAGTGGCAGCGGTTCCTAGTACGTACCAAGTGGACCTACTAC 815
Db	819	CCAATCATGTGTGGGGAGATCAATCTCTTTATCCAAAGCTTACTTTTGTCTATTCG 878

QY	816	CCGGTCATGATCTTCGGCCGAGTCAACCTCTTCATCCAGACCTTTTATTGCTCCTCACC 875
Db	879	AGACGTTACGTTCTTCATGATGAGCCTTGAACATAGCTAGCTGGATCTTGTGTTCTGGACGTGG 938
QY	876	AGCGCGGAGTCCCTGACCGCGCTCTAACTTAATGGTATCGCGGTTCCTGACGCTGG 935
Db	939	TTTCTCTTTTATGATCTCTTCCCTACCAACTGGCAAGAGAGATCATCTTTGCTTTT 998
QY	936	TTCCCGCTCTTGTCTTGTCTCCGAACTGGCCTGAACGGTTCGGGTTCGCTCCTCATC 995
Db	999	AGCATGGCGCTCACGGGATTCAGCAGCTTCAGTCTGTTTAAACCATTTTGGCGGAGAT 1058
QY	996	AGCTTTGGCGTTCAGCGGATTCAGCAGCTTCAGTCTCAGCTCAACCACTTCTCCGGCGAC 1055
Db	1059	GTTTACACGGTCCGCTTAATGGAAACGATGTTTGGAGAACAAACAGCTGGTACGCTT 1118
QY	1056	ACATAGCTGGGCCCCCAAGGCGACACTGGTTCAGAGAGCAGCAAGGAGGACCATC 1115
Db	1119	GATATATCGTAGTTCGTATATGATGTTGTTCTTTGGCGGTTCAGTTTCACTAGAG 1178
QY	1116	GATATCAGTCCCGACCGTGGATGGTCTTTGTTGGGCTGCAGTTCCAGTTGGAG 1175
Db	1179	CATCATCTGTTCCCTAGGTACCGGCTGCTCATCTCAGGGGAGTTTCGCTGTGGTTCAG 1238
QY	1176	CACCACTTCTCCCTAGGTGCGCGTGGCAGCTTAGGAAGATTCGCGCTTGGCTCGG 1235
Db	1239	GAGCTTTGTAAGAAGCATAATCTACCGTATAGGAGTCTTTCGTTGGTGGGA-G--GCTAAT 1295
QY	1236	GACTTGTGAAGAAGCAGCGGATGCGGTATAGGAGTCTCGGTTTGGGACGAGCTAAT 1295
Db	1296	GTGTGACGCTCAGGACTCTGAGGAAAGCGCGGTTCAAGC 1336
QY	1296	GTGAGCAATTCGGACGCTGAGGATCGCGGCTTCAGGC 1336
RESULT	4	
LOCUS	HACYTB5RN	1591 bp mRNA PLN
DEFINITION	H. annuus mRNA for extraplastidial fusion protein.	
ACCESSION	X87143	
NID	91040728	
KEYWORDS	cytochrome B5 domain; fusion protein.	
SOURCE	Helianthus annuus	
ORGANISM	Eukaryota; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Asterales; Asteraceae; Helianthus.	
REFERENCE	1 (bases 1 to 1591)	
AUTHORS	Sperling, P., Schmidt, H. and Heinz, E.	
TITLE	A cytochrome-b5-containing fusion protein similar to plant acyl lipid desaturases	
JOURNAL	Eur. J. Biochem. 232 (3), 798-805 (1995)	
MEDLINE	96028121	
REFERENCE	2 (bases 1 to 1591)	
AUTHORS	Sperling, P.	
TITLE	Direct Submission	
JOURNAL	Submitted (10-MAY-1995) P. Sperling, Univ Hamburg, Inst. fuer Allgemeine Botanik, Ohnhorststr. 18, 22609 Hamburg, FRG	
COMMENT	Related sequences: AJ224160, AJ224161.	
FEATURES	Location/Qualifiers	
Source	1..1591	
	/organism="Helianthus annuus"	
	/cultivar="inbred line HA89, sfi63 /91"	
	/db_xref="taxon:4232"	
	/dev_stage="21 DAF"	
	/tissue_type="cotyledons of developing sunflower fruits"	
	/cell_type="embryonic storage"	
	/clone="PCR-cDNA"	
	1..1591	
	73..1449	
	/note="N-terminal cytochrome b5 domain - desaturase-like enzyme"	
mRNA	1..1591	
CDS	73..1449	

Db	282	A-T--CTTAAAGATTACTCTGTTCTCTGAGGTCTCTAAAGATTATAGGAAGCTGTGTGTTG	338
QY	286	ACTACCTCAAGGACTCTGAAAGTGTGCGAGATCTCTCAAGGACTACCGGAGGCTTTTGAACG	345
Db	339	AGTTTTCTAAATGGTTTGATGACAAAAAAGGTGCATATTATGTTTGCACATTTGTGCT	398
QY	346	AGATGTCGGGTCCGGGATCTTCGAGAAGAAGGCCACCATCATGTGGACGTTTGTGCG	405
Db	399	TTATAGCAATGCTGTTTGGCTATGAGTGTTTATGGGGTTTTGTTTGTGAGGGTGTTTTGG	458
QY	406	GCCTTCGGGTCAATGATGGCGCAATCGTCTACGGCGTCTGGCGTCGGAGTCCGTCGGAG	465
Db	459	TACATTTGTTTCTCGGTGTTTGATGGGGTTCCTTTGGATTACAGATGCGTTGGATTGGAC	518
QY	466	TTCACATGCTCTCGCGCGCACTGCTGGGCTTTCGTGTGGATCCAAAGCCGCGTATGTGGGC	525
Db	519	ATGATGCTGGGCATTATATGTTAGTGTCTGATTCAAGGCTTTAAATGTTTATGGTATTT	578
QY	526	ATGACTCCGGCCATTACAGGTGATGCCAACCCGTGGATACACAGAAATCACGCACTCA	585
Db	579	TTGCTGCAAAATGTTCTTTACGGAATAAGTATTGGTTGGTGGAAATGGAACCAATAATGCAC	638
QY	586	TAGCAGGCAACATCTCAACCGGAATCAGCATCGGTGGTGGAAAGTGGAACCAACAACGCC	645
Db	639	ATCACATTGCCCTGTAATAGCCTTGAATATGACCCCTGATTTACATATATATACCATTCCTTG	698
QY	646	ACCACCTGCCCTGCAACAGCCTTCGACTACGACCCCGACCTCCACACATCCCGGTATTCG	705
Db	699	TTGTGTTCTCCAAAGTTTTTGGTTTCACCTCACTCTCATTTCTATGAGAAAGGTGTGACTT	758
QY	706	CCGTCTCCACCGCACTCTCAACTCCATCACTCCGTCTCTATGGCGGAGTCTGGAAT	765
Db	759	TTGACTCTTTATCAAGATTCTTTGTAAGTATTCAACATTTGACATTTTACCCCTATTATGT	818
QY	766	TCGACGAAAGTGGCAGGGTTCCTAGTACGCTACCAAGCACTGGACCTACTACCGGTCTATGA	825
Db	819	GTGCTGCTAGGCTCAATATGATGTACAATCTCTCATAACTGTTGTTGACCAAGAAATG	878
QY	826	CTTTCGGCGAGTCAACCTCTTCATCCAGAGCTTTTATGCTCCTCACCAGCGCGAGC	885
Db	879	TGTCTATCGAGCTCAGGAACCTCTTGGGATGCCTAGTGTCTCTCAATTTGGTACCCGTTCG	938
QY	886	TCCTTGACCGCGCTCTAAACTTAAATGGGTATCGCGGTTTTCTGGACGTGGTTCGGCTCT	945
Db	939	TTGTTTCTTTGTTGGCTAAATTTGGGGTGAAGAATATGTTTGGTTATTGCAAGTTTATFCAG	998
QY	946	TCGTATCTTTGTTCCCGCACTGGCCTGAACGGTTCGGGTTTCGTCTCTCATCAGCTTTGCGG	1005
Db	999	TGACTGGAAATGAACAAGTTCAAGTTCTCTTGAACCACTTCTCTTCAAGTCTTATATGTTG	1058
QY	1006	TCAGGGGATTCAGACAGCTCCAGTTTCAGCTCAACCACTTCTCCGGGGACACATACGTGG	1065
Db	1059	GAAAGCCTAAAGGGAATAATTGGTTTGAGAACAAACGATGGGACACTTTGACATTTCTTT	1118
QY	1066	GCCCCCAGGGGACAACTGGTTTCGAGAAGCAGACGAAAGGACGATCGATATACGT	1125
Db	1119	GTCTCTTTGGATGGATTGGTTTCATGGTGGATTGCAATTTCCAAATTTGAGCATCATTTGT	1178
QY	1126	GCCACCGTGGATGGACTGGTTCTTTTGGTGGGCTGCAGTTCAGATTGGAGCACCACATTGT	1185
Db	1179	TTCCCAAGATGCTTAGATCAACCTTAGGAAAATCTCGCCCTACGTGATCGAGTTATGCA	1238
QY	1186	TCCCTAGGCTGCCGCTGGCAGCTTAGAAGATTGCGCCCTTGGCTCGGGACTTGTGTA	1245
Db	1239	AGAAACA	1245
QY	1246	AGAAGCA	1252

RESULT	6		PAT	21-APR-1997
LOCUS	I38430	1685 bp	DNA	
DEFINITION	Sequence 4 from patent US 5614393.			
ACCESSION	I38430			

NID	g2084484									
KEYWORDS	Unknown.									
SOURCE	Unknown.									
ORGANISM	Unclassified.									
REFERENCE	1 (bases 1 to 1685)									
AUTHORS	Thomas,F.L., Reddy,A.S., Nuccio,M., Nunberg,A.N. and Freysinet,G.L.									
TITLE	Production of .gamma.-linolenic acid by a .DELTA.6-desaturase									
JOURNAL	Patent: US 5614393-A 4 25-MAR-1997;									
FEATURES	Location/Qualifiers									
source	1..1685									
BASE COUNT	431 a 277 c 357 g 620 t									
ORIGIN	/organism="unknown"									
	Query Match 10.1%; Score 172; DB 22; Length 1685;									
	Best Local Similarity 57.7%; Pred. No. 1.78e-105;									
	Matches 697; Conservative 0; Mismatches 507; Indels 3; Gaps									
Db	42	CAATGGCTGCTCAAAATCAAGAAATACATTACCTCAGATCAACTCAAGAACACACGATAAAC	101							
Qy	46	CAATGGAGGCGAAGCTAAGAGTATATACGCGGGAGGACCTCCCGCCACACAAAGT	105							
Db	102	CCGAGATCTATGATCTCGATTCAAGGAAAGCCATGATTTTCGGATTTGGGTGAAAG	161							
Qy	106	CCGCGGATCTCTGATCTCCATCCAGGCAAGGCTACGACTGCTCTCGGTGGCGCGG	165							
Db	162	ACCATCCAGTGGCAGCTTTCCCTTGAAGAGTCTTGCTGGTCAAGAGGTAAGTATGTCAT	221							
Qy	166	AGCACCCGCGCGGCGAGGTCCCGCTCCTCAGTCTGCGCGCCAGGACGTCACCGACGCCT	225							
Db	222	TTGTTGCATTCCATCTCCCTCTACATGGAAGAATCTTGATAAGTTTTTCACCTGGGTATT	281							
Qy	226	TCATTGGGTACCCCGGCACCGCGTGGCGCATCTGGATCCGCTCTTCACCGGCTACT	285							
Db	282	A-T--CTTAAGATTACTCTGTTCTGAGGTTTCTAAAGATTATAGAAGCTTGTGTTTG	338							
Qy	286	ACTACCTCAAGGACTTCGAAGTGTGCGAGATCTCCAAGGACTACCGGAGCCTTTTGAACG	345							
Db	339	AGTTTTCTAAATGGTTTTGTATGACAAAAAGGTCAATTATGTTTGCAACTTTTGCT	398							
Qy	346	AGATGTCCGGTCCGGGATCTTCGAGAAGAAGGGCCACCACATCATGTGGACGTTGCTCG	405							
Db	399	TTATAGCAATGCTCTTTGCTATGAGTCTTTATGGGTTTTGTTTTGTAGGGTGTTTTGG	458							
Qy	406	CGTTGGCGTCATGATGGCGGCATCTGCTACCGCGGTGTGGCGTCGGAGTCGGTGGAG	455							
Db	459	TACATTTGTTTTCTGGGTGTTTGTATGGGGTTCTTTTGGAFTCAGAGTGGTTGGATTGGAC	518							
Qy	466	TTACATGCTCTGGCGCAGCTGCTGGGCTTGCTGTGGATCCAAAGCCGCTATGTGGGCC	525							
Db	519	ATGATGTGGGCATATATGGTATGTCTGATTCAGGGCTTAATAGTTTATGGGTATTT	578							
Qy	526	ATGACTCCGGCCATTAGCAGGTGATGCAACCCGCTGGATACAAACAGAAATCACCAACTCA	585							
Db	579	TGCTGCAAAATGCTTCTCAGGAATAGTATGGTTGGTGGAAATCGAACCCATAATGCCAC	638							
Qy	586	TAGCAGGCAACATCCTAACCGGAATCAGATCGCGTGGTGGAAAGTGACCCCAACAGCCCC	645							
Db	639	ATCACATGCGCTGTAATAGCCTTGAATATGACGCTGATTTTACAAATATATACCAATTCCTTG	698							
Qy	646	ACCACCTCGCTGCAACAGCCTCGACTACGACCCCGACCTCCAGCACATCCCGGTATTCG	705							
Db	699	TTGTGCTTCCAAAGTTTTTTGGTTTCACTCACCTCTCAATTTCTATGAGAAAAAGGTTGACTT	758							
Qy	706	CCGTCTCCACCCGACTCTTCAACTCCCATCACTCCGCTCTTCTATGGCGAGTCCCTGAAAT	765							
Db	759	TTGACTCTTTATCAAGATCTTTGTAGTTTATCAACATTTGGACATTTTACCCTTATATGT	818							
Qy	766	TCGACGAAGTGGCACGGTTCTTAGTACGTACCGACACTCGGACCTACTACCCGGTATGTA	825							
Db	819	GTGCTGTAGGCTCAATATGTATGACAATCTCTCAATATGTTGTTGACCAAGAGAAATG	878							

RESULT	6
LOCUS	
DEFINITION	
ACCESSION	

QY	826	TCCTGGCGGAGTCAACCTCTTCATCCAGACCTTTTATTGCTCCTCACCAGCGCGACG	885
Db	879	TGTCCTATCGAGCTCAGGAAGTCTTGGGATCGCTAGTGGTCTCGATTTGGTACCGGTGC	938
QY	886	TCCTGACCGCGCTCTAACTTAATGGGTATCGCGGTTTCTGGACGTGGTCCCGCTCT	945
Db	939	TTCTTTCTTTGTTGCTTAATTTGGGTGAAAGATTAATGTTGTTATTGCAAGTTTATCAG	998
QY	946	TCGTATCTTGTCTCCGAAGTCCGCTGAACGGTTCGGGTTCTCGCTCATCAGCTTTGCGG	1005
Db	999	TGACTGGAATGCAACAGTTTCAGTTCTCTTGAACCACTTCTCTTCAAGTGTATTGTTG	1058
QY	1006	TCAGCGGATCCAGCAGCTCCAGTTTCACGCTCAACCACTTCTCCGGGACACATACGTG	1065
Db	1059	GAAGGCTTAAGGGAATAATGTTGTTGAGAAACAACGGATGGGACACTTGACATTTCTT	1118
QY	1066	GCCCCCRAAGGCGCACTGTTGAGAGAGCAGAGGAAGGAGCATCGATATACGT	1125
Db	1119	GTCTCTCTGGATGATGTTTTCATCGGTGATTCGAATTCGAATTCGACATCTTGT	1178
QY	1126	GCCACCGTGGATGGAGTGGTCTTGGTGGGCTGCAGTTCAGTTCGAGTGGAGCACCATTTG	1185
Db	1179	TTCCCAAGATCGCTAGATCAACCTTAGGAAATCTCGCCCTACGTGATCGAGTTATGCA	1238
QY	1186	TCCTAGGCTGCCGCTGGGCACCTTAGGAGATTGCGCCCTTGGCTCGGAGCTTGTA	1245
Db	1239	AGAAACA 1245	
QY	1246	AGAACA 1252	
RESULT	7		
LOCUS	BOUT9010	1687 bp	PLN
DEFINITION	Borago officinalis delta 6 desaturase mRNA, complete cds.		02-MAY-1997
ACCESSION	U79010		
NID	92062402		
KEYWORDS			
SOURCE	Borago officinalis.		
ORGANISM	Borago officinalis		
REFERENCE			
AUTHORS	Sayanova, O., Smith, M.A., Lapinskas, P., Stobart, A.K., Dobson, G., Christie, W.W., Shewry, P.R. and Napier, J.A.		
TITLE	Expression of a borago desaturase cDNA containing an N-terminal cytochrome b5 domain results in the accumulation of high levels of delta6-desaturated fatty acids in transgenic tobacco		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 94 (8), 4211-4216 (1997)		
REFERENCE	97268723		
AUTHORS	Sayanova, O., Smith, M.A., Shewry, P.R. and Napier, J.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-NOV-1996) Cell Biology, IACR-Long Ashton Research Station, Long Ashton, Bristol BS18 9AF, UK		
FEATURES			
Source	1. 1687		
CDS	/organism="Borago officinalis"		
	/db_xref="taxon:13363"		
	42. 1388		
	/function="haem-binding protein"		
	/codon_start=1		
	/product="delta 6 desaturase"		
	/db_xref="PID:92062403"		
	/translation="MAAQIKKVIITDELKNHDKPGLWTSIOGKAYDYSWYKDRHFGS		
	SFPLKSLAGQEVTDFAFHPASTWKNLDKFFTYGLKDYVSSEVSKDYRKIVFFFSK		
	MGLYDKKHIFMTCFLCFVAMFVMSYIGLVFCEGLVHLFSGCLMGLFWIQSGWIGHD		
	AGHYWVSDSLRNKFMGIFAANCLSGISIGWKNHNAHHIACNSLEYPDLOIYIPFL		
	VYSSFEFSLSHFEKRLTFDSLSRFVSYOHWTFFPIMCARLNWYVQSLIMLLTK		
	RNVSTRHELGLCVFSIWIPLVLSCLPNWGERIMFVIASLVTFGNGQQVQFSLNHFS		
	SVYVGPKNWNFEKQTGDTLDISCPPMWNDFHGLQFQIEHLLFPKPRCNLRKISIP		

BASE COUNT	441 a	276 c	356 g	614 t
ORIGIN				
Query Match	10.1%	Score 172;	DB 25;	Length 1687;
Best Local Similarity	57.7%;	Pred. No. 1.78e-105;		
Matches	697;	Conservative 0;	Mismatches 507;	Indels 3; Gaps 2;
Db	40	CAATGGCTGCTCAAAATCAAGAAATACATTTACCTCAGATGAACACTCAAGAACACAGATAAAC	99	
QY	46	CAATGGAGGCGAAGCTTAAGAAGTATATCAGCGGAGGAGCCTCCGCGCCACACACAAGT	105	
Db	100	CCGAGATCTTATGATCTCGATTCAGGGAAGCCTATGATGTTTCGGATGGGTGAAG	159	
QY	106	CCGGCACTCTGGAUCTCCATCCAGGCAAGGCTAGACTGCTCTCGGTGGGCGCGG	165	
Db	160	ACCATCCAGGTGGCAGCTTTCCCTTGAAGAGTCTTGCTGCTCAAGAGGTAACACTGATGAT	219	
QY	166	AGCACCCCGCGGCGAGGTCCCGCTCCTCAGTCTGCGCGCCGACGACGCTACCGAGCGCT	225	
Db	220	TTGTTGATTCATTCCTGCTCTACATGGAAGATCTTGATAAGTCTTCTACTGGGTATT	279	
QY	226	TCATTGGCTACCAACCGGCGACGCGTGGGCGCATCTGGATCCGCTCTTCCACCGCTACT	285	
Db	280	A-T--CTTAAAGATTACTCTGTTCTCAGGTTTCTTAAGATTATAGGAAGCTTGTGTTTG	336	
QY	286	ACTACCTCAAGGACTTCGAAGTGTGGAGATCTCCAAGGACTACCGGAGGCTTTTGAACG	345	
Db	337	AGTTTCTTAAATGGGTTTGTATGACAAAAAGGTCTATATTATTTTGCAACTTTGTGCT	396	
QY	346	AGATGCGCGTCCGGATCTTCGAGAAGAAGGCCACCATCATGTGGACGTTCTGTCG	405	
Db	397	TTATAGCAATGCTGTTGCTATGAGTCTTTATGGGGTTTTGTTTGTGAGGGTGTGTTTG	456	
QY	406	CGGTTGGGTCATGATGGCGCAATCGTCTACGGCGTGTGGCGTCCGAGTCCGTCGGAG	465	
Db	457	TACATTTGTTTCTGGGTGTTGATGGGTTTCTTTGGATTACAGTGGTGGATTGGAC	516	
QY	466	TTACATGCTCTCGGCGCACTGCTGGGCTTGTGTGGATCCCAAGCCGCGTATGTGGGCC	525	
Db	517	ATGATGTCGGGCTTATATGTTAGTGTCTGATTCAAGGCTTAATAAGTTTATGGTATT	576	
QY	526	ATGATCCGGCCATACCAGGTGATGCCAACCCGTTGATACACAGATACAGCAACTCA	585	
Db	577	TTGCTGCAAAATGCTTTTCAGGAATAAGTATTGGTTGGTGAATAATGAACCATATGAC	636	
QY	586	TAGCAGCAACATCTTAACCGGAATCAGATCGCGTGGTGAAGTGGACCCACACAGGCC	645	
Db	637	ATCAGATGCTGTATAGCTTGAATATGACCTGATTTACAAATATATACCATTCCTTG	696	
QY	646	ACCACCTCGCTGCACAGCTCGACTACGACCCCGACCTCCAGCATCCTCCGCTATTTCG	705	
Db	697	TTGTGCTTCCAAGTTTGTGGTTCACCTCCTCATTTCTATGAGAAAGTGTGACTT	756	
QY	706	CCGTCTCCACCCGACTCTTCAACTCCATCAGCTCGGTCTTCTATGGCGAGTCTCGAAT	765	
Db	757	TTGACTCTTTTATCAAGATTTCTTGTAAAGTTATCAACATTTGGACATTTTACCTATTATG	816	
QY	766	TCAGCAAGTGGCAGCTTCTTAGTACGTACCAAGTACGACCTGACCTACTACCGGTGATG	825	
Db	817	GTGCTGCTAGGCTCAATATATGTATCAATCTCTCAATATGTTGTTGACCAAGAGAATG	876	
QY	826	CTTTGGCGGAGTCAACCTCTTATCCAGACCTTTTATTGCTCTCCTACCAAGGCGGACG	885	
Db	877	TGTCCTATCCAGCTCATGAACCTCTTGGGATGCCTAGTGTCTCGATTTGGTACCGGTGC	936	
QY	886	TCCTGACCCGCTCTAAACTTAATGGGTATCGCGGTTTCTTGGACGCTGTTTCCCGCTCT	945	
Db	937	TTGTTCTTTGTTGCTTAATTTGGGTGAAAGAAATATGTTGTTGTTTATTGCAAGTTTGTG	996	
QY	946	TCGTATCTGCTCTCCGAACTGCCCTGAACGGTTCGGGTTGCTGCTCTCATCAGCTTTGCGG	1005	
Db	997	TGACTGGAATGCAACAAGTTTCAGTTCTCTTGAACCACTTCTCTTCAAGTGTGTTTATGTTG	1056	

LOCUS AR024229 965 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5795961.
ACCESSION AR024229
NID 93977523
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 965)
AUTHORS Wallace,T.Paul, Harris,W.J., Carr,F.J., Old,L.J., Welt,S. and Kitamura,K.
TITLE Recombinant human anti-Lewis b antibodies
JOURNAL Patent: US 5795961-A 22 18-AUG-1998;
FEATURES
source Location/Qualifiers
1..965
/organism="unknown"
BASE COUNT 192 a 170 c 226 g 205 t 172 others
ORIGIN

Query Match 2.7%; Score 46; DB 22; Length 965;
Best Local Similarity 17.4%; Pred. No. 7.60e-12;
Matches 26; Conservative 76; Mismatches 44; Indels 3; Gaps 3;

Db 786 VHYSGGVSTCTASDITTSYCWVGRGWGDYGGYTNVNGRGRVTMADTSSNSRSS 845
QY 104 GTCGGCGATCTCTGGATCTCCATCCAGGCAAGTCTACGACTCTCTCGTGGCGGC 163

Db 846 VTAADTAVYCVGRSVDSDGWDYGGTTVTVSSHTVKDMTSSSSASVGDRTVTCRSST 905
QY 164 GGA-GCACCCCGGGCGAGTCCCGTCTCTCAGTCTGGC-CGGC-CAGGACGTACCGCA 220

Db 906 HGMNTYVYKAKYRVSNSRSGVSRS 934
QY 221 CGCTTTCATTCGTACCAACCGGCGCACGG 249

LOCUS RYENRA 281 bp DNA PLN 05-APR-1996
DEFINITION Secale cereale nitrate reductase gene, partial cds.
ACCESSION L40154
NID 9685205
KEYWORDS nitrate reductase.
SOURCE Secale cereale DNA.
ORGANISM Secale cereale
Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Liliopsida; Poales; Poaceae; Secale.
REFERENCE 1 (bases 1 to 281)
AUTHORS Zhou,J., Kilian,A., Warner,R.L. and Kleinhofs,A.
TITLE Variation of nitrate reductase genes in selected grass species
JOURNAL Genome 38 (5), 919-927 (1995)
MEDLINE 96076701
FEATURES
source Location/Qualifiers
1..281
/organism="Secale cereale"
/db_xref="taxon:4550"
CDS
1..>281
/codon_start=1
/product="nitrate reductase"
/db_xref="PID:g685206"
/translation="DDDKQFTMSVRKHSKSAWIVVHGVDYDCTGFLKDHDPGGADS
ILINAGTCTEEFDATHSDKAKALLDAHRIGELITTGAGYNSDNSVHAAP"

BASE COUNT 60 a 101 c 75 g 45 t
ORIGIN

Query Match 2.5%; Score 43; DB 24; Length 281;
Best Local Similarity 63.0%; Pred. No. 6.13e-10;
Matches 104; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Db 61 TGGATCTCTCCACGACGATGACGCTTCTCAGCCGCTTCTCAGGACACACCCCGGT 120
QY 117 TGGATCTCTCCACGACGATGACGCTTCTCAGCTGCTCTCGTGGCGGCGGACACCCCGGC 176

LOCUS 8
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
NID 92724471
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner,F., Scheiflinger,F. and Falkner,F.Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
source Location/Qualifiers
1..7218
/organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 5.3%; Score 91; DB 22; Length 7218;
Best Local Similarity 0.3%; Pred. No. 2.20e-43;
Matches 1; Conservative 229; Mismatches 139; Indels 0; Gaps 0;

Db 1067 YY 1126
QY 634 CCCACAAGCGCCACCACTCGCTGCAACAGCTCGACTACGACCGCGACCTCCAGCACA 693

Db 1127 YY 1186
QY 694 TCCCGTATTTCGGCTCCACCGGACTCTCAACTCCATCCATCGCTCTTCTATGGCC 753

Db 1187 YY 1246
QY 754 GAGTCTGAATTCGAGAGTGGCAGGTCTCTAGTACGACTACCACTGACCTACT 813

Db 1247 YY 1306
QY 814 ACCGGTCATGATCTCGGCGAGTCAACCTCTCATCCAGACTTTTATGCTCTCA 873

Db 1307 YY 1366
QY 874 CCAGGCGCAGCTCCCTGACCGGCTCTAACTTAATGGGTATCGCGGTTTCTGGACGT 933

Db 1367 YY 1426
QY 934 GGTTCGGCTCTTCGATCTTGTCTCCGAACTGGCGTGAACGGTTCGGGTCTGCTCTCA 993

Db 1427 YYYYYYTG 1435
QY 994 TCAGCTTG 1002

RESULT 9

[illegible]

[illegible]

repeat_region	3707. .3728 /note="(A)22" /rpt_type=tandem /rpt_unit=A 4366. .4661 /rpt_family="Alu" 5327. .5602 /rpt_family="Alu" 6586. .6956 /rpt_family="L1" 6647. .6684 /note="(CA)19" /rpt_type=tandem /rpt_unit=CA 7113. .7373 /rpt_family="Alu" complement(7830. .8185) /standard_name="possible repeat" 8258. .8503 /rpt_family="Alu" 9070. .9387 /rpt_family="Alu" complement(9740. .9845) /rpt_family="MER42" complement(10440. .11015) /rpt_family="Alu" 11950. .12350 /rpt_family="Alu" 12057. .12085 /note="(A)29" /rpt_type=tandem /rpt_unit=A 12365. .12645 /rpt_family="Alu" 13727. .13750 /note="(AC)12" /rpt_type=tandem /rpt_unit=AC 13783. .14024 /rpt_family="L1" 14175. .14470 /rpt_family="Alu" complement(14906. .15259) /standard_name="possible repeat" 15300. .15613 /rpt_family="Alu" 16671. .16690 /note="(T)20" /rpt_type=tandem /rpt_unit=T complement(16675. .16977) /rpt_family="Alu" complement(16993. .17085) /rpt_family="MER42" complement(17678. .18276) /rpt_family="Alu" 19305. .19583 /rpt_family="Alu" 19914. .19945 /note="(GTTT)8" /rpt_type=tandem /rpt_unit=GTTT complement(19943. .20222) /rpt_family="Alu" 21202. .21496 /rpt_family="Alu" 21736. .22035 /rpt_family="Alu" 22017. .22038 /note="(A)22" /rpt_type=tandem /rpt_unit=A 22321. .22457
---------------	---

```
repeat_region /rpt_family="Alu"  
repeat_region complement(22638..22981)  
repeat_region /standard_name="possible repeat"  
repeat_region 23022..23326  
repeat_region /rpt_family="Alu"  
repeat_region 23473..23761  
repeat_region /rpt_family="Alu"  
repeat_region 23744..23767  
repeat_region /note="(A)24"  
repeat_region /rpt_type=tandem  
repeat_region /rpt_unit=A  
repeat_region complement(24539..24694)  
repeat_region /rpt_family="MER42"  
repeat_region complement(25349..25713)  
repeat_region /rpt_family="Alu"  
repeat_region complement(25727..26471)  
repeat_region /rpt_family="Alu"  
repeat_region 27191..27477  
repeat_region /rpt_family="Alu"  
repeat_region 27774..28057  
repeat_region /rpt_family="Alu"  
repeat_region 28040..28066  
repeat_region /note="(A)27"  
repeat_region /rpt_type=tandem  
repeat_region /rpt_unit=A  
repeat_region 28712..28930  
repeat_region /rpt_family="MER20"  
repeat_region complement(28769..28838)  
repeat_region /note="GRAIL 2 excellent exon, frame 0"  
repeat_region 28987..29214  
repeat_region /rpt_family="Alu"  
repeat_region join(29485..29559,29598..29779)  
repeat_region /note="95% identity dbSTS:G14522 (SHGC-11312)"  
repeat_region 29495..29976  
repeat_region /standard_name="AI027942"  
repeat_region /note="100% identity EST ov84a10.x1"  
repeat_region complement(30401..30536)  
repeat_region /note="GRAIL 2 excellent exon, frame 2"  
repeat_region complement(30682..30733)  
repeat_region /rpt_family="MIR"  
repeat_region complement(31573..31724)  
repeat_region /note="GRAIL 2 excellent exon, frame 1"  
repeat_region complement(32159..32232)  
repeat_region /note="GRAIL 2 excellent exon, frame 2"  
repeat_region 32388..32488  
repeat_region /rpt_family="MLT1"  
repeat_region 32617..32908  
repeat_region /rpt_family="Alu"  
repeat_region 32977..33088  
repeat_region /rpt_family="MLT1"  
repeat_region complement(33670..33785)  
repeat_region /rpt_family="Alu"  
repeat_region complement(34021..34144)  
repeat_region /note="GRAIL 2 excellent exon, frame 2"  
repeat_region complement(35238..35331)  
repeat_region /note="GRAIL 2 excellent exon, frame 1"  
repeat_region 36392..36663  
repeat_region /rpt_family="Alu"  
repeat_region 36901..37222  
repeat_region /note="100% identity EST ou55c09.x1"  
repeat_region /db_xref="dbEST:AI025011"  
repeat_region 36901..37164  
repeat_region /standard_name="STSG-9983"  
repeat_region /db_xref="dbSTS:G26554"  
repeat_region complement(37595..37654)  
repeat_region /note="GRAIL 2 excellent exon, frame 0"  
repeat_region complement(join(38063..38218,38462..38578,38741..38995,  
39071..39205,39532..39630,39935..40048,40300..40410,  
40503..40561,41868..41972,42103..42225,42492..42569,  
44379..44507))  
repeat_region /standard_name="histidyl-tRNA synthetase"  
repeat_region /note="78% 100% protein identity GenPept:U18937"  
repeat_region complement(38069..38215)
```

```
misc_feature /note="GRAIL 2 excellent exon, frame 0"  
misc_feature complement(38462..38578)  
...  
Note: remainder of annotations omitted.  
Query Match 2.2%; Score 38; DB 27; Length 74371;  
Best Local Similarity 13.3%; Pred.No. 7.36e-07;  
Matches 28; Conservative 101; Mismatches 79; Indels 2; Gaps 2;  
Db 15925 SKSCMSRSKSRGWYRSWKYRCAMWMTCKSSKWCWSYRMRMKYSCSYCYS-SGK 15983  
QY 774 GTGGCAGGTTCCCTAGT-CAGCTACCAGCAGCTGGACCTACTACCGGGTCATGATCTTCGG 832  
Db 15984 KYWRCRSMYWTTCYYSKYYSMSYCTCTSGWRWWSKGRSWMYASRSGCSCSMMCM 16043  
QY 833 CCGAGTCAACCTCTTCATCCAGACCTTTTATTGCTCCTCACCAGGCGGACGTCCTGA 892  
Db 16044 CRCSCMSMKMWWTTTITTKRTTWTWTKRKGASASRGKSKRCMSYGGKSKTKMKMTC 16103  
QY 893 CCGCGCTCTAAACTTAATGGGTATCGCGGTTTCTGGACGTGGTTCGCGCTCTTCGTATC 952  
Db 16104 YYGMYSCSYRKMYSRYCCGCGYSSSYCKS 16133  
QY 953 TTGCTCCCGAAGTGGCTGAACGGTTCGG 982
```

Search completed: Thu Feb 18 16:22:21 1999
Job time : 6092 secs.

W P E F L (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 17:40:33 1999; Maspar time 244.06 Seconds
948.786 Million cell updates/sec

Linear output not generated.

Title: >US-08-934-254-26
Description: (1-1702) from US08934254.seq
Perfect Score: 1702

N.A. Sequence: 1 CCCCAAAATTTTCATGTT.....TTTGTGTAATAAAAAA 1702
Comp: GGGGTTTTTAAAGTAACAA.....AAAAACCATTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: n-geneseq32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 9.489; Variance 5.740; scale 1.653

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	172	10.1	1685	21	T30395	2.48e-94
2	47	2.8	204	1	N81164	8.70e-13
3	43	2.5	91	9	Q51746	1.79e-10
4	43	2.5	204	1	N81164	1.79e-10
5	40	2.4	91	9	Q51746	9.00e-09
6	36	2.1	114	12	Q70469	1.47e-06
7	36	2.1	114	12	Q70468	1.47e-06
8	36	2.1	114	12	Q70465	1.47e-06
9	36	2.1	114	12	Q70470	1.47e-06
10	36	2.1	114	12	Q70467	1.47e-06
11	35	2.1	114	12	Q70472	5.12e-06
12	36	2.1	114	12	Q70465	1.47e-06
13	36	2.1	114	12	Q70468	1.47e-06

C 14	36	2.1	114	12	Q70467	Generic DNA sequence	1.47e-06
C 15	35	2.1	114	12	Q70469	Generic DNA sequence	5.12e-06
C 16	34	2.0	114	12	Q70466	Generic DNA sequence	1.77e-05
C 17	34	2.0	114	12	Q70471	Generic DNA sequence	1.77e-05
C 18	32	1.9	114	12	Q70473	Generic DNA sequence	2.02e-04
C 19	32	1.9	114	12	Q70472	Generic DNA sequence	2.02e-04
C 20	32	1.9	114	12	Q70466	Generic DNA sequence	2.02e-04
C 21	32	1.9	114	12	Q70470	Generic DNA sequence	2.02e-04
C 22	32	1.9	114	12	Q70465	Generic DNA sequence	2.02e-04
C 23	30	1.8	172	32	T76363	Human endothelin-1 an	2.19e-03
C 24	30	1.8	172	32	T76363	Human endothelin-1 an	2.19e-03
C 25	30	1.8	172	32	T76363	Human endothelin-1 an	2.19e-03
C 26	31	1.8	540	32	T75438	Mortierella alpina cy	6.70e-04
C 27	30	1.8	565	6	Q35072	HCV envelope region n	2.19e-03
C 28	29	1.7	172	32	T76363	Human interleukin 8 a	7.04e-03
C 29	29	1.7	190	32	T76452	Chymase antisense oli	7.04e-03
C 30	29	1.7	190	32	T76452	Chymase antisense oli	7.04e-03
C 31	27	1.6	39	7	Q51787	Mixed oligonucleotide	6.94e-02
C 32	27	1.6	67	24	T14322	Primer used in the la	6.94e-02
C 33	28	1.6	114	12	Q70471	Generic DNA sequence	2.23e-02
C 34	28	1.6	114	12	Q70473	Generic DNA sequence	2.23e-02
C 35	28	1.6	130	32	T76152	Human vascular cell a	2.23e-02
C 36	27	1.6	168	32	T76270	Human MDNCF antisense	6.94e-02
C 37	26	1.5	90	21	T13616	ME#1 random peptide 1	2.12e-01
C 38	26	1.5	201	2	N70195	Streptomyces protease	2.12e-01
C 39	26	1.5	201	2	N70194	Signal portion of gen	2.12e-01
C 40	26	1.5	1558	22	T28255	Survival motor neuron	2.12e-01
C 41	26	1.5	1560	22	T18828	Human survival motor	2.12e-01
C 42	26	1.5	1582	22	T18831	Human survival motor	2.12e-01
C 43	26	1.5	1582	22	T28259	Survival motor neuron	2.12e-01
C 44	26	1.5	3871	2	N71302	HSV-1 gb and surround	2.12e-01
C 45	26	1.5	5852	2	Q11710	Dictyostellium plasmid	2.12e-01

ALIGNMENTS

RESULT 1

ID T30395 standard; DNA; 1685 BP.

AC T30395;

DE 15-SEP-1996 (first entry)

DE Borage delta-6-desaturase gene.

KW Delta-6-desaturase; gamma-linolenic acid; transgenic plant; borage;

KW polyunsaturated fatty acid; octadecatetraonic acid;

KW chilling resistance; oilseed; ss; ds.

OS Borage officinalis.

PH Key Location/Qualifiers

FT cds 44..1390

FT /*tag= a

PN W09621022-A2.

PD 11-JUL-1996.

PF 28-DEC-1995; IB1167.

PR 30-DEC-1994; US-366779.

PA (RHON) RHONE POULENC AGROCHIMIE.

PI Freyssinet GL, Nuccio M, Nunberg AN, Reddy AS, Thomas TL;

DR WPI: 96-333997/33.

DR P-PSDB: R98455.

PT Transgenic plants comprising the borage delta-6-desaturase gene

PT show increased production of gamma linolenic acid and having

PT increased resistance to chilling

PS Claim 2; Page 51-52; 75pp; English.

CC A DNA clone (T30395) codes for borage delta-6-desaturase (R98455),

CC which catalyses the conversion of linoleic acid to gamma-linolenic

CC acid (GLA). It was isolated from a borage membrane-bound polysomal

CC library using probes based on abundantly expressed seed storage

CC protein cDNAs and with an isolated partial cDNA clone. The gene

CC can be incorporated into a vector, pref. incorporating a

CC tissue-specific promoter, for the expression of delta-6-desaturase

CC in transgenic plants, esp. sunflower, soybean, maize, tobacco,

CC peanut, carrot or oilseed rape, resulting in increased GLA prodn.

CC Alteration of plant lipids may also lead to improved chilling

CC resistance.

SQ Sequence 1685 BP; 431 A; 277 C; 357 G; 620 T;

Query Match 10.1%; Score 172; DB 21; Length 1685;
 Best Local Similarity 57.7%; Pred. No. 2,48e-94;
 Matches 697; Conservative 0; Mismatches 507; Indels 3; Gaps 2;

Db 42 caatgctgctcaaatcaagaatacattaccctcagatgaactcaagaaccacgataaac 101
 QY 46 CAATGGAGGGCGAAGCTAAGAAGTATATCACGGCGGAGGACCTCCGCCGCCACAACAAGT 105
 Db 102 ccggagatctatggtatctcagatcaaggaaagcctatgatcttccgattgggtgaaag 161
 QY 106 CCGGCGATCTCTGGATCTCCATCCAGGGCAAGGTCTACGACTGCTCTCGGTGGCGCGCG 165
 Db 162 accatccaggtgagcgtcttccttgaaagctcgtcgtcaagaggttaactgatgat 221
 QY 166 AGCACCCCGCGCGGAGGCTCCGCTCTCAGTCTGCGCGCGGAGGACGTCACCGAGCCT 225
 Db 222 ttgttgatccatcctgctcctacatggaagaactcttgataagtttttcaactgggtatt 281
 QY 226 TCATTGCGGTACACCCCGGCGACGGGTGCGGCATCTGATCGGCTCTTCACCGGCTACT 285
 Db 282 a-t--cttaagatctactctgttctgaggtttctaaagattatggaagcttgggttg 338
 QY 286 ACTACTCAAGGACTTCGAAGTGTGGAGATCTCAAGGACTACCGGAGCTTTTGAAGC 345
 Db 339 agttttctaaaaatggtttgtatgacaaaaggtcattatattgtttgcaactttgtct 398
 QY 346 AGATGTCGGGTCCGGGATCTTCGAGAAAGAGGGCCACCACATCATGTGGAGCTTCGTCG 405
 Db 399 ttatgcaatgctgtttgtatgagtttgggtttgtttgtgaggggtgttttg 458
 QY 406 GGGTTGCGGTATGATGGGGCAATCGTCTACGGGTGCTGCGTCCGAGTCGCGAG 465
 Db 459 tacattgtttctcgttgggtttgtatgaggtttcttggattcagagtggttggattgg 518
 QY 466 TTCACATGCTCTGCGCGGACTGCTGGGCTTCTGTGGATCCAAAGCGCGTATGTGGGCC 525
 Db 519 atgactgctgggcatatgtagtctgattcaggtcatttaaggtttatgggtattt 578
 QY 526 ATGATCCGGCGCATACCAAGGTGATGCCAACCCGCTGGATFACACAGAAATCACGCCAAC 585
 Db 579 ttgctgcaattgttttcaggaataagattgttgggtggaatggaaccataatgac 638
 QY 586 TAGCAGGCAACATCTAACCGGAATCAGATCGCTGTTGGAAGTGGACCCACACAGGCC 645
 Db 639 atcattgctctgtaagccttgatgacccctgatttacaatatatacattccttg 698
 QY 646 ACCACTCGCTCGAACAGCCTCGACTACGACCCCGACCTCCAGCACATCCCGCTATTCG 705
 Db 699 ttgtgttccaagtgttttgggttcaactcactcattctcatttatgagaaaggttgact 758
 QY 706 CGGTCTCCACCGGACTCTTCAACTCCATCACCTCGGTCTCTATGCGCGAGTCTGAAAT 765
 Db 759 ttgactctttatcaagattcttgaattatcaacattgacattttacccattattgt 818
 QY 766 TCGACGAAGTGCACGGTTCCTAGTACGTACACAGCTGGACCTTACCCGGGTATGA 825
 Db 819 gtgctgtaggtcctaatatgtatgataatctctcattatgttggtagcagaagaaatg 878
 QY 826 TCTTGGCGGAGTCAACCTTTCATCCAGACCTTTTATTTGCTCTCACCGGCGGAGC 885
 Db 879 tgtcctatcgagctcaggaactcttgggtgctgtagtctcatttggtagcccggtgc 938
 QY 886 TCCCTGACCGGCTCTAACTTAATGGGTATCGCGGTTTCTGGACGTGTTCTCCGCTCT 945
 Db 939 ttgtttcttctgctaatatgggtgaaagaaatattgtttgttatttgaagtttatcag 998
 QY 946 TCGTATCTGTCTCCGAACCTGGCTGAACCGGTTCGGGTCTCTCTCATCAGCTTTGCGG 1005
 Db 999 tgactggaatgcaacaagtctcagttccttgtaaccactctcttcaagtgtttatgtg 1058
 QY 1006 TCAGCGCATCGACGATCGCTCAGCTTCAGCTCAACCACTTCTCCGGCGACATACGTGG 1065
 Db 1059. gaaagcctaaaggggaaataatgtgttgagaaacaaaggtgggacatttgacatttct 1118

QY 1066 GCCCCCCCAAGGGCGACACTGGTTCGAGACGACGAGGAAGGACGATACACGT 1125
 Db 1119 gtcctctctggtgattgattgtttcattcattggtgattgcaattccaaattgagcatctt 1178
 QY 1126 GCCCACCGTGGATGGACTGTTCTTTGGTGGCTGAGTTCAGTTGGAGCACCATTTGT 1185
 Db 1179 ttcccaagatgctcagatgcaacaccttagaataatctcgcctcagtgatgagattatgca 1238
 QY 1186 TCCCTAGGCTCGCGGTGGCAGCTTAGGAAGATTCGCGCCCTTGGCTCGGAGCTTGTGA 1245
 Db 1239 aqaaca 1245
 QY 1246 AGAAGCA 1252

RESULT 2
 ID N81164 standard; DNA; 204 BP.
 AC N81164;
 DT 08-NOV-1990 (first entry)
 DE Base substituted E.coli beta-galactosidase alpha-fragment.
 KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT misc_feature 19..69
 FT /tag= a
 FT /function= multiple cloning site
 FT primer_bind 187..204
 FT /tag= b
 PN EP-285123-A.
 PD 05-MAY-1988; 105163.
 PF 30-MAR-1988; 105163.
 PR 03-APR-1987; US-034819.
 PA (SUSO) SUOMEN SOKERI OY.
 PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
 DR WPI: 88-279927/40.
 PT Introducing random point mutations into nucleic acids -
 PT by prep of single stranded template, annealing a primer, elongation,
 PT misincorporation, completion of molecules and screening.
 PS Disclosure: P: English.
 CC Random point mutations were introduced into the alpha fragment of
 CC E.coli beta-galactosidase. The wild type sequence was obtained as a
 CC single stranded template and an oligonucleotide was hybridised to
 CC it to generate a popn of DNA molecules which terminate at all
 CC possible nucleotide positions within a specified region. The
 CC variable 3' ends generated in this way are used as primers for
 CC reverse transcriptase. Nucleotides are misincorporated by the
 CC transcriptase and the molecules are completed to forms that can be
 CC amplified and then expressed in a suitable host-vector system.
 CC The sequence covers all 176 diff base substitutions, most of which
 CC occurred singularly in any given mutant.
 CC See also P80575.
 SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 2.8%; Score 47; DB 1; Length 204;
 Best Local Similarity 12.8%; Pred. No. 8.70e-13;
 Matches 15; Conservative 60; Mismatches 42; Indels 0; Gaps 0;
 Db 75 yycdhvccgcyymrtthhrrmrbnvdydnrsdaaayccyrsvkdyccynachhdd 134
 Cp 447 CCAGCACCGCTAGACGATTCGCCCATCATGACCGCAACGCCGACGACGTCACATGA 388
 Db 135 byyvybbvynvhnncnccbnhvhvbnhnnrwnvayvthdarrrdvhccvchc 191
 Cp 387 TGTGTGGCCCTCTCTCGAAGATCCCGGACCGGACATCTCTGTTCAAAGCCTCC 331

RESULT 3
 ID Q51746 standard; cDNA; 91 BP.
 AC Q51746;
 DT 31-MAY-1994 (first entry)
 DE Oligonucleotide probe MK14-A
 KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;

```

QY 524 TGGAAAGTGGACCCACAACAGCCGCCACCTCGCTCGCAACAGCCTCGACTAGCAGCCCGAC 68
Db 94 yrmrbnvdyrdnsdaawccyyrriyskvdyccynachddhdyvvybbvavhnhnncnc 153
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 684 CTCAGCACATCCCGGTATTCCGGCTTCACCGGACTCTTCAACTCCATCAGCTCGGTC 74
Db 154 ccbnhvhcnvbnhbnhrnwyrzhdarrddvh 185
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 744 TTCTATGGCGAGTCGTCGAAATTCGACGAAGT 775

RESULT 5
ID Q51746 standard; cDNA; 91 BP.
AC Q51746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PF 24-MAY-1993; 108325.
PR 26-MAY-1992; US-889651.
PR (BECT ) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI: 93-378844/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 2.4%; Score 40; DB 9; Length 91;
Best Local Similarity 5.5%; Pred. No. 9,00e-09;
Matches 3; Conservative 47; Mismatches 4; Indels 1; Gaps

Db 10 gssvhsyvvvhhvshhshvhhvhhvsvvvvhhvvhv-vhhvhyhvyvsvtc 63
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1537 GGGAAACGACACCGACTAGCAAAGTTAAGCAAAGTTAAGCTGACGTAGTGCTTC 1483

RESULT 6
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PR (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain

```

PS Disclosure: Page 35; 255pp; English.
 CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
 CC This generic formula can be represented as follows: X(TGC)(NNB)10-
 CC (TGC)(NNB)5Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
 CC sites (X is not the same as Y) that are not specified further. This
 CC sequence generates peptides that are cloverleaf in structure. Other
 CC generic sequences are shown in Q70463-68. Other specific peptides
 CC generated by these generic sequences are shown in R65150-54. TSARs are
 CC concatenated heterofunctional proteins or peptides, comprising at least
 CC two functional regions - a binding domain with affinity for a ligand and
 CC a second effector peptide portion that is chemically or biologically
 CC active. They may further comprise a linker peptide between the 2 domains.
 CC The oligonucleotides are also designed so that the expressed peptide
 CC contains 2 or 4 cysteine residues positioned in, or flanking, the
 CC unpredicted or variant residues. These residues confer some degree of
 CC conformational rigidity to the peptides. The TSARs or compans. comprising
 CC a TSAR binding domain can be used in vivo to deliver a chemically or
 CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
 CC or enzyme, to the specific target or on the cell. They can also replace
 CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
 CC and therefore circumvent the need for complex methods of hybridoma
 CC formation or in vivo antibody production. The TSARs are easily
 CC characterised and have designed activity allowing direct and rapid
 CC detection in a screening process.
 CC Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

```

Query Match      2.1%; Score 36; DB 12; Length 114;
Best Local Similarity 5.5%; Pred. NO. 1.47e-06;
Matches 6; Conservative 32; Mismatches 72; Indels 0; Gaps 0;

Db      2  gccnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 61
||      : : : : : : : : : : : : : : : : : : : : : :
Qy      91  GCGCCACACAAGTCGGCGAATCTGGATCTCCATCCAGGCGCAAGTCTACGACTGCT 150

Db      62  nbnbnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 111
||      : : : : : : : : : : : : : : : : : : : : : :
Qy      151  CTCGGTGGCGGGAGACACCCGGGGGGAGGTCGCGTCTCTCACTCG 200

```

RESULT	7	
ID	Q70458	standard; DNA; 114 BP.
DE	Q70458;	
DT	05-APR-1995	(first entry)
DE	Generic DNA sequence to generate a random TSAR peptide library	
DE	TSAR; totally synthetic affinity reagent; synthetic; binding domain;	
KW	effector domain; concatenated heterofunctional protein; linker;	
KW	direct; rapid; detection; screening; treatment; generic; ss.	
	Synthetic.	
Key		
	Location/Qualifiers	
FT	misc_feature	55..60
FT		/*tag= a
FT		/note= "this sequence represents 'z'; z can be a
FT		sequence of 6, 9 or 12 nucleotides (see
FT		comments)."
FT		

PN WO9418318-A.
PD 18-AUG-1994. PF U00977.
PR 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
PI WPI: 94-279739/34.
DR P-PGDB: R65154.
DR PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35: 255pp; English.
CC CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)101. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.

Other specific peptides generated by these generic sequences are shown in R5151-54. TSARs are concatenated heterofunctional proteins or peptides, comprising at least two functional regions - a binding domain with affinity for a ligand and a second effector peptide portion that is chemically or biologically active. They may further comprise a linker peptide between the 2 domains. The oligonucleotides are also designed so that the expressed peptide contains 2 or 4 cysteine residues positioned in, or flanking, the unpredicted or variant residues. These residues confer some degree of conformational rigidity to the peptides. The TSARs or compns., comprising a TSAR binding domain can be used *in vivo* to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or *in vivo* antibody production. The TSARs are easily characterised and have designed activity allowing direct and rapid detection in a screening process.

Sequence 114 Bp: 2 C; 2 G; 2 T;
 0 A; 2 C;

Query Match	2.1%	Score 36;	DB 12;	Length 114;
Best Local Similarity	3.6%;	Pred. No. 1.47e-06;		
Matches	4;	Conservative	Mismatches 74;	Gaps 0;

Dbb	3 bnnbnbnbnbnbnbnbnbnbnbnbtgcannbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn	62
Qy	: : : : : : : : : : : : : :	:
62 TAAGAAGTATACGCGGGAGGACCCTCCGCCGCACCAAGTCGCGCATCTTGAT		121
Dbb	63 bnnbnbnbnbnbnbnbnbtgcannbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn	114
Qy	: : : : : : : : : : : :	:
122 CTCATCAGGGGAAGGTCTACGACTGCTCTCGGTGGCGCGGAGCACCCC		173

RESULT	8	
ID	Q70465	standard; DNA; 114 BP.
AC	Q70465;	
DE	05-APR-1995	(first entry)
DT	DE	Generic DNA sequence to generate a random TSAR peptide library.
DE	TSAR;	totally synthetic affinity reagent; synthetic; binding domain
KW	TSAR;	effector domain; concatenated heterofunctional protein; linker;
KW	direct;	rapid; detection; screening; treatment; generic; ss.
KW	Synthetic.	
OS		
FH	Key	Location/Qualifiers
FT	misc_feature	55..60
FT		/tag= a
FT		/note= "this sequence represents '2'; 2 can be a
FT		sequence of 6, 9 or 12 nucleotides (see
FT		comments)"
FT		

WO9418318-A.

PD	18-AUG-1994.
PF	01-FEB-1994;
PR	01-FEB-1993;
PR	30-DEC-1993;
PR	31-JAN-1994;
PA	(UYNC-) UNIV
PI	Fowlkes DM,
DR	WPI; 94-2797

P-PSDB; R65150 and R65151.
PT Identifying proteins or peptides(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)5(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC Q70469-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned

[illegible]

RESULT	10	
ID	Q70467	standard; DNA; 114 BP.
AC	Q70467;	
DT	05-APR-1995	(first entry)
DE	Generic DNA sequence to generate a random TSAR peptide library.	
KE	TSAR: totally synthetic affinity reagent; synthetic; binding domain;	
KW	effector domain; concatenated heterofunctional protein; linker;	
KW	direct; rapid; detection; screening; treatment; generic; ss.	
OS	Synthetic.	
SV		
FH	Key	Location/Qualifiers

```

FT      /tag= a
FT      /note= this sequence represents 'Z'; Z can be a
FT      sequence of 6, 9 or 12 nucleotides (see
FT      comments)"
FT      PN
PN      WO9418318-A.
PD      18-AUG-1994.
PD      01-FEB-1994; U000977.
PR      01-FEB-1993; US-013416.
PR      30-DEC-1993; US-176500.
PR      31-JAN-1994; US-189331.
PR      (UYNC-) UNIV NORTH CAROLINA.
PA      Fowlkes DM, Kay BK;
PI      WPI: 94-279739/34.
DR      P-PSDB: R65153.
DR

```

[illegible]

QY 71 TATCAGGCGGAGGACCTCCGCCGCCACAACAAGTCCGGGATCTCTGGATCTCCATCCA 130

[illegible]


```

sequence of 6, 9 or 12 nucleotides (see
comments)"
PT WO9418318-A.
PN 18-AUG-1994.
PD 01-FEB-1994; U00977.
PF 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PS 31-JAN-1994; US-189331.
PT (UINC-) UNIV NORTH CAROLINA.
PA Fowlkes DM, Kay BK;
PI P-PSDB; R65154.
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)14. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 36; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 1.47e-06;
Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
Cp 659 GCAGCGGAGTGGGCGTGTGGTCCACTCCACCGCGATGCTGATCCGGTTAG 600
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
599 GATGTCCTGCTATGATGATGCGGTGATCTGTTGATCCACGGGTTGGCATC 548

RESULT 14
ID Q70467 standard; DNA; 114 BP.
AC Q70467;
DE 05-APR-1995 (first entry)
KW Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PS 30-DEC-1993; US-176500.
PT (UINC-) UNIV NORTH CAROLINA.
PA Fowlkes DM, Kay BK;
PI P-PSDB; R65154.
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)7(TGC)(NNB)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 36; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 1.47e-06;
Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
Cp 659 GCAGCGGAGTGGGCGTGTGGTCCACTCCACCGCGATGCTGATCCGGTTAG 600
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
599 GATGTCCTGCTATGATGATGCGGTGATCTGTTGATCCACGGGTTGGCATC 548

RESULT 14
ID Q70467 standard; DNA; 114 BP.
AC Q70467;
DE 05-APR-1995 (first entry)
KW Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PS 30-DEC-1993; US-176500.
PT (UINC-) UNIV NORTH CAROLINA.
PA Fowlkes DM, Kay BK;
PI P-PSDB; R65154.
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)14. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 36; DB 12; Length 114;
Best Local Similarity 2.8%; Pred. No. 1.47e-06;
Matches 3; Conservative 33; Mismatches 70; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
Cp 737 GGTGATGAGTTGAAGAGTCGGTGGAGACGGCGAATACGGGATGCTGGAGTCGGG 678
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 108
Cp 677 GTCGTAGTCGAGGCTGTTGCAGCGAGGTGTCGGCGTGTGGGTC 632

RESULT 15
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DE 07-APR-1995 (first entry)
KW Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PS 30-DEC-1993; US-176500.
PT (UINC-) UNIV NORTH CAROLINA.
PA Fowlkes DM, Kay BK;
PI P-PSDB; R65154.
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)14. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

```

Search completed: Thu Feb 18 17:49:40 1999
Job time : 547 secs.

[W][A][T][E][R][M][A][N]
[A][L][G][O][R][I][T][H][M]
[T][M]

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 11:36:06 1999; MasPar time 13.22 Seconds
917.710 Million cell updates/sec

Par output not generated.

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pep
Perfect Score: 3515
Sequence: 1 MEGEAKYITAEEDLRHNKS.....LNSAPCPKKGYGAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swissprot

Statistics: Mean 51.005; Variance 92.089; scale 0.554

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Pred. No.	
Result	Score	Query Match	Description	ID	
1	220	6.3	INDUCIBLE NITRATE REDU	3.17e-20	
2	219	6.2	NITRATE REDUCTASE 1 (E	4.87e-20	
3	217	6.2	NITRATE REDUCTASE 1 (E	1.15e-19	
4	215	6.1	NITRATE REDUCTASE, CLO	2.72e-19	
5	212	6.0	NITRATE REDUCTASE, CLO	9.79e-19	
6	209	5.9	NITRATE REDUCTASE, CLO	3.51e-18	
7	203	5.8	CYTOCHROME B5.	4.45e-17	
8	203	5.8	CYTOCHROME B5.	4.45e-17	
9	203	5.8	CYTOCHROME B5.	4.45e-17	
10	205	5.8	NITRATE REDUCTASE 2 (E	1.91e-17	
11	202	5.7	NITRATE REDUCTASE 2 (E	6.78e-17	
12	200	5.7	NITRATE REDUCTASE (EC	1.57e-16	
13	202	5.7	NITRATE REDUCTASE (EC	6.78e-17	
14	199	5.7	NITRATE REDUCTASE (EC	2.39e-16	
15	196	5.6	NITRATE REDUCTASE (NAD	8.35e-16	
16	193	5.5	CYTOCHROME B5.	2.90e-15	
17	195	5.5	CYTOCHROME B5. SEED IS	1.27e-15	
18	195	5.5	INDUCIBLE NITRATE REDU	1.27e-15	
19	194	5.5	NITRATE REDUCTASE 1 (E	1.92e-15	
20	193	5.5	NITRATE REDUCTASE 2 (E	2.90e-15	
21	193	5.5	NITRATE REDUCTASE (EC	2.90e-15	
22	192	5.5	NITRATE REDUCTASE (NAD	4.39e-15	
23	191	5.4	CYTOCHROME B5, OUTER M	6.64e-15	

ALIGNMENTS

RESULT 1				STANDARD;		PRT;		890 AA.	
ID	NIA2_SOYBN								
AC	P39870;								
DT	01-FEB-1995 (REL. 31, CREATED)								
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)								
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)								
DE	INDUCIBLE NITRATE REDUCTASE 2 (EC 1.6.6.1) (NR).								
GN	INR2.								
OS	GLYCINE MAX (SOYBEAN).								
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;								
OC	FABACEAE.								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-CV. WILLIAMS;								
RA	WU S., LU Q., KRIZ A.L., HARPER J.E.;								
RL	PLANT MOL. BIOL. 29:491-506(1995).								
CC	-!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST								
CC	STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.								
CC	-!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.								
CC	-!- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)								
CC	AND ONE MOLYBDENUM ATOM.								
CC	-!- SUBUNIT: HOMODIMER (BY SIMILARITY).								
CC	-!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE								
CC	N-TERMINAL DOMAIN.								
CC	-!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.								
CC	-!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE								
CC	C-TERMINAL DOMAIN.								
EMBL	U13987; G537627; ..								
DR	PROSITE; PS00191; CYTOCHROME_B5; 1.								
DR	PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.								
DR	OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;								
KW	NITRATE ASSIMILATION; MULTIGENE FAMILY.								
FT	METAL 169 169 MOLYBDENUM-PTERIN (POTENTIAL).								
FT	METAL 223 223 MOLYBDENUM-PTERIN (POTENTIAL).								
FT	DISULFID 408 408 INTERCHAIN (POTENTIAL).								
FT	BINDING 552 552 HEME LIGAND (BY SIMILARITY).								
FT	BINDING 575 575 HEME LIGAND (BY SIMILARITY).								
SQ	SEQUENCE 890 AA; 99945 MW; 9EBDDAA9 CRC32;								
Query Match 6.3%; Score 220; DB 1; Length 890;									
Best Local Similarity 41.1%; Pred. No. 3.17e-20;									
Matches 30; Conservative 16; Mismatches 26; Indels 1; Gaps 1;									
Db	513	MNTSKTYTMSR-VRRHNSADSANIIVGHVYDCRFLKDHDPGGDSILINAGTCTEEF 571							
Qy	1	MEGEAKYITAEEDLRHNKSGDLWLSIOGVYDCSRWAHEFGGVEVPLLSLAGQDVDAF 60							

```

Db 572 EAIHSDKAKQMLE 584
| | | | | | | | | |
QY 61 IAYHPGTAWRHLD 73

RESULT 2
ID NIAL_PHAVU STANDARD; PRT; 881 AA.
AC P39865;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE 1 (EC 1.6.6.1) (NR-1).
GN NIAL OR NR1.
OS PHASEOLUS VULGARIS (KIDNEY BEAN) (FRENCH BEAN).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
OC FABACEAE.
RN [1]
RP SEQUENCE FROM N.A.
STRAIN-CV. SAXA; TISSUE-SHOOT;
HOFF T., STUMMANN B.M., HENNINGSSEN K.W.;
PHYSIOL. PLANTARUM 82:197-204(1991).
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -!- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)
CC AND ONE MOLYBDENUM ATOM.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
DR EMBL; X53603; G21019;
DR PIR; S25445; S25445.
DR PROSITE; PS00191; CYTOCHROME B5; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
KW NITRATE ASSIMILATION; MULTIGENE FAMILY.
FT METAL 167 167 MOLYBDENUM-PTERIN (POTENTIAL).
FT METAL 221 221 MOLYBDENUM-PTERIN (POTENTIAL).
FT DISULFID 406 406 INTERCHAIN (POTENTIAL).
FT BINDING 550 550 HEME LIGAND (BY SIMILARITY).
FT BINDING 573 573 HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 881 AA; 99222 MW; C5DA985D CRC32;

Query Match 6.2%; Score 219; DB 1; Length 881;
Best Local Similarity 41.1%; Pred.No. 4.87e-20;
Matches 30; Conservative 15; Mismatches 27; Indels 1; Gaps 1;

Db 511 MNTATKSY-SLSEVRHNRDMSAWIIVGHVYDCTFLKDHGPGDSILLNAGTDCTEEF 569
| | | | | | | | | |
QY 1 MEGEAKKVITAEDLRHKNKSGDLWISIQKVYDCSRWAAEHGPGVPLLSLAGQVTDFAF 60
| | | | | | | | | |

Db 570 EAIHSDKAKQMLE 582
| | | | | | | | | |
QY 61 IAYHPGTAWRHLD 73

RESULT 3
ID NIAL_ARATH STANDARD; PRT; 917 AA.
AC P11832;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE 1 (EC 1.6.6.1) (NRI).
GN NIAL.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC CAPRALES; CRUCIFERAE.
RN [1]
RP SEQUENCE FROM N.A.; AND HERBICIDE RESISTANCE.
STRAIN-CV. COLUMBIA;

```

```

RX MEDLINE: 93287999.
RA WILKINSON J.Q., CRANFORD N.M.;
RL MOL. GEN. 239:289-297(1993).
RN [2]
RP SEQUENCE OF 342-360 AND 525-917 FROM N.A.
RX MEDLINE: 89091069.
RA CHENG C., DENDNEY J., NAM H., DEN BOER B.G.W., GOODMAN H.M.;
RL EMO J. 7:3309-3314(1988).
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -!- COFACTOR: EACH SUBUNIT OF THE ENZYME CONTAINS 1 EQUIVALENT OF FAD,
CC HEME IRON, AND MOLYBDENUM-PTERIN AS PROSTHETIC GROUPS. THE HEME
CC GROUP IS CALLED CYTOCHROME B-557.
CC -!- SUBUNIT: HOMODIMER.
CC -!- TISSUE SPECIFICITY: ROOT, LEAF, AND SHOOT.
CC -!- WHEN MUTATED CONFERS RESISTANCE TO THE HERBICIDE CHLORATE.
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
DR EMBL; Z19050; G22757;
DR EMBL; X13434; G930001;
DR EMBL; X13436; G295791;
DR PIR; S01640; S01640.
DR PIR; S35228; S35228.
DR PIR; S32018; S32018.
DR HSSP; P17571; 2CND.
DR PROSITE; PS00191; CYTOCHROME B5; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
KW NITRATE ASSIMILATION; MULTIGENE FAMILY; HERBICIDE RESISTANCE.
FT METAL 197 197 MOLYBDENUM-PTERIN (POTENTIAL).
FT METAL 251 251 MOLYBDENUM-PTERIN (POTENTIAL).
FT DISULFID 436 436 INTERCHAIN (POTENTIAL).
FT BINDING 580 580 HEME LIGAND (BY SIMILARITY).
FT BINDING 603 603 HEME LIGAND (BY SIMILARITY).
FT VARIANT 198 198 A -> T (LOSS OF ACTIVITY).
SQ SEQUENCE 917 AA; 103125 MW; 599233D3 CRC32;

Query Match 6.2%; Score 217; DB 1; Length 917;
Best Local Similarity 38.4%; Pred.No. 1.15e-19;
Matches 28; Conservative 18; Mismatches 26; Indels 1; Gaps 1;

Db 541 MNTASKMYSISE-VRKHTADSAWIIVGHVYDCTFLKDHGPGDTSILNAGTDCTEEF 599
| | | | | | | | | |
QY 1 MEGEAKKVITAEDLRHKNKSGDLWISIQKVYDCSRWAAEHGPGVPLLSLAGQVTDFAF 60
| | | | | | | | | |

Db 600 EAIHSDKAKQMLE 612
| | | | | | | | | |
QY 61 IAYHPGTAWRHLD 73
| | | | | | | | | |

RESULT 4
ID NIAL_BRANA STANDARD; PRT; 911 AA.
AC P39867;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE, CLONE PNBRI405 (EC 1.6.6.1) (NR).
GN NIAL.
OS BRASSICA NAPUS (RAPE).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC CAPRALES; CRUCIFERAE.
RN [1]
RP SEQUENCE FROM N.A.
STRAIN-CV. LISANDRA;
RX MEDLINE: 96222419.
RA FUKUOKA H., OGAWA T., MINAMI H., YANO H., OHKAWA Y.;
RL PLANT PHYSIOL. 111:39-47(1996).
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.

```

CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -!- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)
CC AND ONE MOLYBDENUM ATOM.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
DR EMBL: D38219; G540485; -
DR PROSITE; PS00191; CYTOCHROME_B5; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
DR OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
KW NITRATE ASSIMILATION; MULTIGENE FAMILY.
FT METAL 191 191 MOLYBDENUM-PTERIN (POTENTIAL).
FT METAL 245 245 MOLYBDENUM-PTERIN (POTENTIAL).
FT DISULFID 430 430 INTERCHAIN (POTENTIAL).
FT BINDING 574 574 HEME LIGAND (BY SIMILARITY).
FT BINDING 597 597 HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 911 AA; 102252 MW; 57A3E33B CRC32;

Query Match 6.1%; Score 215; DB 1; Length 911;
Best Local Similarity 38.4%; Pred. No. 2.72e-19;
Matches 28; Conservative 18; Mismatches 26; Indels 1; Gaps 1;

Db 535 MNTAKMYSME-VKRNKSANESAIIVHGHYDCTRLKDHGGSDSLINAGTDCITEEF 593
QY 1 MEGEAKYITAEEDLRNRKNSGDLWISIQGVKDYDCSRWAAEHPPGVEVPLLSLAGQDVDAF 60

Db 594 EAIHSDRAKALLE 606
QY 61 IAYHPGTAWRHLD 73

RESULT 5
ID NTA2_BRANA STANDARD; PRT; 911 AA.
AC P39868;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE, CLONE PBNER1412 (EC 1.6.6.1) (NR).
GN NTA2.
OS BRASSICA NAPUS (RAPE).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC CAPPARALES; CRUCIFERAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. LISANDRA;
RC MEDLINE; 96222419.
RA YUKUOKA H., OGAWA T., MINAMI H., YANO H., OHKAWA Y.;
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -!- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)
CC AND ONE MOLYBDENUM ATOM.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
DR EMBL: D38220; G540487; -
DR PROSITE; PS00191; CYTOCHROME_B5; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
KW NITRATE ASSIMILATION; MULTIGENE FAMILY.
FT METAL 191 191 MOLYBDENUM-PTERIN (POTENTIAL).
FT METAL 245 245 MOLYBDENUM-PTERIN (POTENTIAL).
FT DISULFID 430 430 INTERCHAIN (POTENTIAL).
FT BINDING 574 574 HEME LIGAND (BY SIMILARITY).
FT BINDING 597 597 HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 911 AA; 102200 MW; F0CFEF43 CRC32;

Query Match 6.0%; Score 212; DB 1; Length 911;
Best Local Similarity 39.7%; Pred. No. 9.79e-19;
Matches 29; Conservative 16; Mismatches 27; Indels 1; Gaps 1;

Db 535 MNTAKMYSME-VKRNKSANESAIIVHGHYDCTRLKDHGGSDSLINAGTDCITEEF 593
QY 1 MEGEAKYITAEEDLRNRKNSGDLWISIQGVKDYDCSRWAAEHPPGVEVPLLSLAGQDVDAF 60

Db 594 EAIHSDRAKALLE 606
QY 61 IAYHPGTAWRHLD 73

RESULT 6
ID CYB5_TOBAC STANDARD; PRT; 136 AA.
AC P49098;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CYTOCHROME B5.
OS NICOTIANA TABACUM (COMMON TOBACCO).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC SOLANALES; SOLANACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LEAF;
RC MEDLINE; 94325476.
RA SMITH M.A., STOBART A.K., SHEWRY P.R., NAPIER J.A.;
CC -!- FUNCTION: CYTOCHROME B5 IS A MEMBRANE BOUND HEMOPROTEIN WHICH
CC FUNCTION AS AN ELECTRON CARRIER FOR SEVERAL MEMBRANE BOUND
CC OXYGENASES. MAY PLAY A KEY ROLE IN THE MODIFICATION BY
CC DESATURATION OF FATTY ACIDS IN THE ENDOPLASMIC RETICULUM, WHICH IN
CC THE DEVELOPING SEED IS UTILIZED FOR MEMBRANE SYNTHESIS AND IN THE
CC DEVELOPMENTALLY REGULATED PRODUCTION OF LARGE AMOUNTS OF STORAGE
CC LIPIDS. IS INVOLVED IN THE REDUCTION OF CYTOCHROME P-450 AND
CC MAY THEREFORE BE INVOLVED IN FLAVONOID BIOSYNTHESIS IN THE PETALS.
CC -!- SUBCELLULAR LOCATION: MICROSMAL MEMBRANE. BOUND TO THE
CC CYTOPLASMIC SIDE OF THE ENDOPLASMIC RETICULUM (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: IS HIGHLY EXPRESSED IN DEVELOPING SEEDS.
CC MODERATELY EXPRESSED IN FLOWERS, AND IS EXPRESSED AT LOW
CC LEVELS IN THE LEAF.
CC EMBL: X71441; G296386; ALT_INIT.
DR EMBL: X68140; G19853; -
DR PROSITE; PS00191; CYTOCHROME_B5; 1.
KW ELECTRON TRANSPORT; TRANSMEMBRANE; HEME; MICROSOME; MULTIGENE FAMILY.
FT BINDING 40 40 HEME LIGAND (BY SIMILARITY).
FT BINDING 64 64 HEME LIGAND (BY SIMILARITY).
FT TRANSMEM 107 127 POTENTIAL.
FT CONFLICT 10 11 LA -> EF (IN G19853).
FT CONFLICT 105 105 MISSING (IN G19853).
SQ SEQUENCE 136 AA; 14979 MW; 6F34BDD9 CRC32;

Query Match 5.9%; Score 209; DB 1; Length 136;
Best Local Similarity 41.8%; Pred. No. 3.51e-18;
Matches 33; Conservative 17; Mismatches 27; Indels 2; Gaps 2;

Db 1 MGGETKVFTLAE-VSQHNNAKDCWLIVSGKYVDYTKFLDDHPPGGDEVLLSATGKDATDFF 59
QY 1 MEGEAKYITAEEDLRNRKNSGDLWISIQGVKDYDCSRWAAEHPPGVEVPLLSLAGQDVDAF 60

Db 60 EDVGHSSSARAMLDEYVVG 78
QY 61 IAY-HPGTAWRHLDPLETG 78

RESULT 7
ID CYB5_YEAST STANDARD; PRT; 120 AA.
AC P40312;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE CYTOCHROME B5.
GN CYB5 OR YNL111C OR N1949.
OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMICOTINA; HEMIASCOMICETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F1100;
RX MEDLINE: 94237477.
RA TRUAN G., EPINAT J.-C., ROUGEULLE C., CULLIN C., POMPON D.;
RL GENE 142:123-127(1994).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97245296.
RA DE ANTONI A., D'ANGELO M., DAL PERO F., SARTORELLO F., PANDOLFO D.,
RA PALLAVICINI A., LANFRANCHI G., VALLE G.;
RL YEAST 13:261-266(1997).
CC -!- FUNCTION: CYTOCHROME B5 IS A MEMBRANE BOUND HEMOPROTEIN WHICH
FUNCTION AS AN ELECTRON CARRIER FOR SEVERAL MEMBRANE BOUND
OXYGENASES. IT PLAYS A ROLE IN FATTY-ACID DESATURATION AND IS ALSO
INVOLVED IN SEVERAL STEPS OF THE STEROL BIOSYNTHESIS PATHWAY,
PARTICULARLY IN THE 4-DEMETHYLATION OF THE 4,4'-DIMETHYL
ZYMOSTEROL.
CC -!- SUBCELLULAR LOCATION: MICROSOmal MEMBRANE. BOUND TO THE
CYTOPLASMIC SIDE OF THE ENDOPLASMIC RETICULUM (BY SIMILARITY).
DR EMBL: L22494; G431762; -;
DR EMBL: Z69382; E221832; -;
DR EMBL: Z71387; E239931; -;
DR PIR: S47919; S47919.
DR SGD: L0002498; CYB5.
DR PROSITE: PS00191; CYTOCHROME_B5; 1.
KW ELECTRON TRANSPORT; TRANSMEMBRANE; HEME; MICROsome.
FT BINDING 37 37 HEME LIGAND (BY SIMILARITY).
FT BINDING 61 61 HEME LIGAND (BY SIMILARITY).
FT TRANSMEM 98 118 POTENTIAL.
FT CONFLICT 17 17 E -> Q (IN REF. 1).
SQ SEQUENCE 120 AA; 13297 MW; 9BEA894F CRC32;

Query Match 5.8%; Score 203; DB 1; Length 120;
Best Local Similarity 38.7%; Pred. No. 4.45e-17;
Matches 29; Conservative 19; Mismatches 25; Indels 2; Gaps 2;

Db 2 PKVY-SYOEVAHGPENFWIIIDDKVYDVQFQKDEHPGGDEIINDLGQDATESFVDIG 60
Qy 5 AKKIYTDLRHKNKSGDLWISIOGKVYDCSRWAAEHPPGVEPLLSLAGQDVTDAFIAY- 63
Db 61 HSDEALRLKGLIYG 75
64 HPGTAWRHLDPLFTG 78

RESULT 8
ID CYB5_ORYSA STANDARD; PRT; 137 AA.
AC P49100;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CYTOCHROME B5.
OS ORYZA SATIVA (RICE).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;
OC CYPERALES; GRAMINEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CALLUS;
RX MEDLINE: 94325476.
RA SMITH M.A., STOBART A.K., SHEWRY P.R., NAPIER J.A.;
RL PLANT MOL. BIOL. 25:527-537(1994).
CC -!- FUNCTION: CYTOCHROME B5 IS A MEMBRANE BOUND HEMOPROTEIN WHICH
FUNCTION AS AN ELECTRON CARRIER FOR SEVERAL MEMBRANE BOUND
OXYGENASES.
CC -!- SUBCELLULAR LOCATION: MICROSOmal MEMBRANE. BOUND TO THE
CYTOPLASMIC SIDE OF THE ENDOPLASMIC RETICULUM (BY SIMILARITY).
DR EMBL: X75670; G414705; -;
DR PROSITE: PS00191; CYTOCHROME_B5; 1.

KW ELECTRON TRANSPORT; TRANSMEMBRANE; HEME; MICROsome.
FT BINDING 41 41 HEME LIGAND (BY SIMILARITY).
FT BINDING 65 65 HEME LIGAND (BY SIMILARITY).
FT TRANSMEM 108 128 POTENTIAL.
SQ SEQUENCE 137 AA; 15296 MW; B204EA8B CRC32;

Query Match 5.8%; Score 203; DB 1; Length 137;
Best Local Similarity 40.5%; Pred. No. 4.45e-17;
Matches 32; Conservative 13; Mismatches 33; Indels 1; Gaps 1;

Db 1 MSNDNKKVYTLLEEVAKHNSKDCWLIIGKVYVSKFLEDHFGDDVLLSSTGKDATDDF 60
Qy 1 MBGEAKKVTIATEDLRHKNKSGDLWISIOGKVYDCSRWAAEHPPGVEPLLSLAGQDVTDAF 60
Db 61 EDVGHITTTARAMMDEYYVG 79
Qy 61 IAY-HPGTAWRHLDPLFTG 78

RESULT 9
ID NIA_LOTJA STANDARD; PRT; 900 AA.
AC P39869;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE (EC 1.6.6.1) (NR).
GN NIA.
OS LOTUS JAPONICUS.
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
OC FABACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. GIFU / B-129;
RA WATERHOUSE R.N., SMYTH A.J., PROSSER I.M., FORDE B.G., CLARKSON D.T.;
RL SUBMITTED (JUL-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -!- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)
AND ONE MOLYBDENUM ATOM.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
N-TERMINAL DOMAIN.
CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
C-TERMINAL DOMAIN.
CC EMBL: X80670; G517356; -;
DR PIR: S47029; S47029
DR PROSITE: PS00191; CYTOCHROME_B5; 1.
DR OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
KW NITRATE ASSIMILATION.
FT METAL 172 172 MOLYBDENUM-PTERIN (POTENTIAL).
FT METAL 226 226 MOLYBDENUM-PTERIN (POTENTIAL).
FT DISULFID 411 411 INTERCHAIN (POTENTIAL).
FT BINDING 556 556 HEME LIGAND (BY SIMILARITY).
FT BINDING 579 579 HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 900 AA; 101420 MW; 2F667418 CRC32;

Query Match 5.8%; Score 203; DB 1; Length 900;
Best Local Similarity 37.0%; Pred. No. 4.45e-17;
Matches 27; Conservative 19; Mismatches 26; Indels 1; Gaps 1;

Db 517 MNTFTKMY-SLSEVKHNSPDSAIIVHGVYDCTRELKDHPPGADSLINAGTDCTEEF 575
Qy 1 MEGEAKKVTIATEDLRHKNKSGDLWISIOGKVYDCSRWAAEHPPGVEPLLSLAGQDVTDAF 60
Db 576 EAIHSDRAKKMLE 588
Qy 61 IAYHPGTAWRHLD 73

RESULT 10

```

ID NIA2_ARATH STANDARD; PRT; 917 AA.
AC P11035;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE 2 (EC 1.6.6.1) (NR2).
GN NIA2 OR CHL3.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC CAPPALES; CRUCIFERAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE; 88276888.
RA CRAWFORD N.M., SMITH M., BELLISSIMO D., DAVIS R.W.;
RL PROC. NATL. ACAD. SCI. U.S.A. 85:5006-5010(1988).
RN [2]
RP SEQUENCE OF 522-917 FROM N.A.
RX MEDLINE; 89091069.
RA CHENG C., DEWDNEY J., NAM H., DEN BOER B.G.W., GOODMAN H.M.;
RL EMBO J. 7:3309-3314(1988).
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE; 93005711.
RA WILKINSON J.O., CRAWFORD N.M.;
RL PLANT CELL 3:461-471(1991).
RN [4]
RP HERBICIDE RESISTANCE.
RX MEDLINE; 93287999.
RA WILKINSON J.O., CRAWFORD N.M.;
RL MOL. GEN. GENET. 239:289-297(1993).
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -!- COFACTOR: EACH SUBUNIT OF THE ENZYME CONTAINS 1 EQUIVALENT OF FAD,
CC HEME IRON, AND MOLYBDENUM-PTERIN AS PROSTHETIC GROUPS. THE HEME
CC GROUP IS CALLED CYTOCHROME B-557.
CC -!- SUBUNIT: HOMODIMER.
CC -!- TISSUE SPECIFICITY: ROOT, LEAF, AND SHOOT.
CC -!- WHEN MUTATED CONFERS RESISTANCE TO THE HERBICIDE CHLORATE.
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
DR EMBL; J03240; G166782; -
DR EMBL; X13435; G930002; -
DR EMBL; S45385; -; NOT_ANNOTATED_CDS.
DR PIR; A31821; RDMUNH.
DR PIR; S01641; S01641.
DR HSSP; P17571; 2CND.
DR PROSITE; PS00191; CYTOCHROME_B5; 1.
KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
KW NITRATE ASSIMILATION; MULTIGENE FAMILY; HERBICIDE RESISTANCE.
FT METAL 191 191 MOLYBDENUM-PTERIN (POTENTIAL).
FT METAL 245 245 MOLYBDENUM-PTERIN (POTENTIAL).
FT DISULFID 433 433 INTERCHAIN (POTENTIAL).
FT BINDING 577 577 HEME LIGAND (BY SIMILARITY).
FT BINDING 600 600 HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 917 AA; 102844 MW; 95608AFB CRC32;

Query Match 5.8%; Score 205; DB 1; Length 917;
Best Local Similarity 38.4%; Pred. No. 1.91e-17;
Matches 28; Conservative 18; Mismatches 26; Indels 1; Gaps 1;

Db 538 MNTAKMSSE-VKKNHSDSADSVIIVGHYDCTRFKLDHPGGSDSLINAGTDCTEEF 596
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1 MEGEAKKVTAEEDLRHNRKSGDLWISIOGVYDCSRWAAEHGPGVPLLSLAGQDVTDAF 60

Db 597 EAIHSDKAKKMLE 609
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 61 IAYHPGTAWRHLD 73

```

```

RESULT 11
ID NIA2_PHAVU STANDARD; PRT; 890 AA.
AC P39866;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE 2 (EC 1.6.6.1) (NR-2).
GN NIA2 OR NR2.
OS PHASEOLUS VULGARIS (KIDNEY BEAN) (FRENCH BEAN).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
OC FABACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SAXA;
RA JENSEN P.E., HOFF T., STUMMANN B.M., HENNINGSEN K.W.;
RL SUBMITTED (AUG-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -!- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)
CC AND ONE MOLYBDENUM ATOM.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.
CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
DR EMBL; U01029; G392992; -
DR PROSITE; PS00191; CYTOCHROME_B5; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
KW NITRATE ASSIMILATION; MULTIGENE FAMILY.
FT METAL 165 165 MOLYBDENUM-PTERIN (POTENTIAL).
FT METAL 219 219 MOLYBDENUM-PTERIN (POTENTIAL).
FT DISULFID 404 404 INTERCHAIN (POTENTIAL).
FT BINDING 548 548 HEME LIGAND (BY SIMILARITY).
FT BINDING 571 571 HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 890 AA; 99995 MW; D952FE9D CRC32;

Query Match 5.7%; Score 202; DB 1; Length 890;
Best Local Similarity 34.2%; Pred. No. 6.78e-17;
Matches 25; Conservative 21; Mismatches 26; Indels 1; Gaps 1;

Db 509 MNTASKMF-SVSEVKKHSPPDSAWIIVGHYVDCRFKLDHPGGTDSILINAGTDCTEEF 567
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1 MEGEAKKVTAEEDLRHNRKSGDLWISIOGVYDCSRWAAEHGPGVPLLSLAGQDVTDAF 60

Db 568 DAHSDKAKKMLE 580
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 61 IAYHPGTAWRHLD 73

RESULT 12
ID NIA_LYCES STANDARD; PRT; 911 AA.
AC P17570;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE (EC 1.6.6.1) (NR).
GN NIA.
OS LYCOPERSICON ESCULENTUM (TOMATO).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC SOLANALES; SOLANACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MANAPAL; TISSUE=LEAF;
RX MEDLINE; 90185211.
RA DANIEL-VEDELE F., DORBE M.F., CABOCHE M., ROUZE P.;
RL GENE 85:371-380(1989).
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.

```


Db 583 DAIHSDKAKKMLE 595

WORLDH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Feb 18 11:36:54 1999; MasPar time 25.99 Seconds
865.947 Million cell updates/sec
[redacted] ar output not generated.

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pep
Perfect Score: 3515
Sequence: 1 MEGAKKYITAEDLRHNKS.....LNSAPCPKKGGEAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrnbl6

1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 49.519; Variance 96.434; scale 0.514

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2433	69.2	458	10	CYTCHROME B5 CONTAINI	0.00e+00
2	2133	60.7	448	10	DELTA 6 DESATURASE.	0.00e+00
3	431	12.3	444	4	BC269730.2	1.11e-58
4	407	11.6	443	5	DELTA6-FATTY-ACID-DESA	6.23e-54
5	405	11.5	473	5	W08D2.4	1.54e-53
6	317	9.0	454	5	T13F2.1.	1.70e-36
7	220	6.3	875	10	NITRATE REDUCTASE (FRA	1.05e-18
8	207	5.9	134	10	CYTCHROME B5 (FRAGMENT	1.91e-16
9	200	5.7	911	10	NADH NITRATE REDUCTASE	3.04e-15
10	200	5.7	911	10	NADH NITRATE REDUCTASE	3.04e-15
11	194	5.5	113	10	CYTCHROME B5 (FRAGMEN	3.17e-14
12	194	5.5	134	10	PUTATIVE CYTOCHROME B5	3.17e-14
13	193	5.5	427	2	HYPOTHETICAL 48.4 KD P	4.68e-14
14	193	5.5	915	10	NITRATE REDUCTASE APOE	2.20e-13
15	189	5.4	368	2	DELTA 6 DESATURASE.	2.20e-13
16	190	5.4	891	3	NITRATE REDUCTASE (NAD	1.50e-13
17	188	5.3	74	10	NITRATE REDUCTASE HEME	3.24e-13
18	187	5.3	146	4	CYTCHROME B5 (FRAGMEN	4.76e-13
19	185	5.3	497	10	NITRATE REDUCTASE (EC	1.03e-12
20	187	5.3	501	10	NITRATE REDUCTASE (EC	4.76e-13

21	185	5.3	629	10	P92920	NITRATE REDUCTASE (EC	1.03e-12
22	183	5.2	94	10	Q39966	NITRATE REDUCTASE (FRA	2.21e-12
23	182	5.2	95	10	Q39993	NITRATE REDUCTASE (FRA	3.23e-12
24	182	5.2	95	10	Q39976	NITRATE REDUCTASE (FRA	3.23e-12
25	182	5.2	95	10	Q39984	NITRATE REDUCTASE (FRA	3.23e-12
26	183	5.2	132	10	O24651	CYTCHROME B5 (FRAGMEN	2.21e-12
27	183	5.2	132	10	O04354	CYTCHROME B5.	2.21e-12
28	184	5.2	487	10	P92921	NITRATE REDUCTASE (EC	1.51e-12
29	181	5.1	94	10	Q41318	NITRATE REDUCTASE (FRA	4.73e-12
30	176	5.0	91	10	Q38791	NITRATE REDUCTASE (FRA	3.14e-11
31	177	5.0	95	10	Q39971	NITRATE REDUCTASE (FRA	2.15e-11
32	177	5.0	95	10	Q39992	NITRATE REDUCTASE (FRA	2.15e-11
33	177	5.0	95	10	Q41613	NITRATE REDUCTASE (FRA	2.15e-11
34	176	5.0	907	3	Q92237	NITRATE REDUCTASE.	3.14e-11
35	174	5.0	907	3	Q00303	NITRATE REDUCTASE.	6.66e-11
36	171	4.9	98	6	Q28726	SOLUBLE CYTOCHROME B5.	2.04e-10
37	167	4.8	171	11	P70116	CYTCHROME B5.	9.02e-10
38	169	4.8	892	3	O13486	NITRATE REDUCTASE.	4.30e-10
39	163	4.6	100	11	Q35768	CYTCHROME B5.	3.93e-09
40	163	4.6	121	10	O22704	PUTATIVE CYTOCHROME B5	3.93e-09
41	160	4.5	877	10	Q42497	NITRATE REDUCTASE.	1.17e-08
42	157	4.5	864	3	O01696	NITRATE REDUCTASE.	3.48e-08
43	153	4.4	112	5	O17091	CYTCHROME B5.	1.46e-07
44	153	4.4	565	3	O13510	L-MANDELATE DEHYDROGEN	1.46e-07
45	151	4.3	138	5	Q93315	C3E10.7.	2.97e-07

ALIGNMENTS

RESULT	ID	PRELIMINARY;	PRT;	458 AA.
AC	Q43469			
DT	01-NOV-1996	(TREMBREL. 01, CREATED)		
DT	01-NOV-1996	(TREMBREL. 01, LAST SEQUENCE UPDATE)		
DT	01-JUN-1998	(TREMBREL. 06, LAST ANNOTATION UPDATE)		
DE	CYTCHROME B5 CONTAINING FUSION PROTEIN.			
OS	HELIANTHUS ANNUUS (COMMON SUNFLOWER).			
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;			
OC	ASTERALES; COMPOSITAE.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. INBRED LINE HA89, SFI63 /91;			
RC	TISSUE=COTYLEDONS OF DEVELOPING SUNFLOWER FRUITS;			
RX	MEDLINE; 96028121.			
RA	SPELING P.; SCHMIDT H.; HEINZ E.;			
RL	EUR. J. BIOCHEM. 232:798-805(1995).			
DR	EMBL; X87143; G1040729; -.			
DR	PFAM; PF00173; heme-1.			
SQ	SEQUENCE 458 AA; 52231 MW; FA2AE27B CRC32;			

Query Match	69.2%;	Score 2433;	DB 10;	Length 458;
Best Local Similarity	66.7%;	Pred. No. 0.00e+00;		
Matches	300;	Conservative 92;	Mismatches 54;	Indels 4;
Gaps	4;			
Db	13	ADGKYITKELKKHNPDLWISILGKVVNTEWAKEHPGDAPLNLACQDVTDAFIA	72	
QY	3	GEAKYITAEDLRHNKSGDLWISIQKVDYDCSWRAEHPGGEVPLLSLAQDVTDAFIA	62	
Db	73	FHPGTANKHLDKFTGYH-LKDYQVSDISDRYKLASEFAKAGMFEKKGHGVYSLCFVS	131	
QY	63	YHPGTANKHLDKFTGYH-LKDYQVSDISDRYKLASEFAKAGMFEKKGHGVYSLCFVS	122	
Db	132	LLLSACVYGVLYSGFWIHLMSGAILGLAWQIAYLGHGDAGHYOMATRGWNKAGFIFIG	191	
QY	123	VMMAIYGVGLASESVGHMLCGALLGLLWIAQYVGHDSGHYQVMPTRGNRTIQLIAG	182	
Db	192	NCITGISIAKWKTHNAHHAACNSLDYDPLQHLPLMLAVSSKLFNSITSVYGRQLTDP	251	
QY	183	NILTGISIAKWKTHNAHHAACNSLDYDPLQHLPLMLAVSSKLFNSITSVYGRQLTDP	242	
Db	252	LARFFVSYQHYLYYPIMCVARVNLQTLILLISKRIPDRGLNLGLTFIETWTFPLVLS	311	
QY	243	VARELVSYQHYLYYPIMCVARVNLQTLILLISKRIPDRGLNLGLTFIETWTFPLVLS	302	

```

AC OC0427;
AD 01-AUG-1998 (TREMBREL. 07, CREATED)
DE DT 01-AUG-1998 (TREMBREL. 07, LAST SEQUENCE UPDATE)
DE DT 01-AUG-1998 (TREMBREL. 07, LAST ANNOTATION UPDATE)
DE BC269730.2.
OS OS
OS HOMO SAPIENS (HUMAN)
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RN SEQUENCE FROM N.A.
RA LAMERDIN J.E., MCCREARY P.M., COLEMAN M., SKOWRONSKI E., ADAMSON A.W.
RA BURKHART-SCHULTZ K., GORDON L., KYLE A., RAMIREZ M., STILLWAGEN S.,
RA PHAN H., VELASCO N., DO L., REGALA W., TERRY A., GARNES J., DANGANAN
RA POUNDSTONE P., CHRISTENSEN M., GEORGESCU A., AVILA J., LIU S., ARTIX
RA ANDREISE T., FRANKHEIM M., AMICO-KELLER G., COFIELD J., DUARTE S.,
RA LUCAS S., BRUCE R., THOMAS P., QUAN G., KRONMILLER B., ARELLANO A.,
RA MONTGOMERY M., OW D., NOLAN M., TRONG S., KOBAYASHI A., OLSEN A.S.,
RA CARRANO A.V.;
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AC004770; G3169158; -.
SQ SEQUENCE 444 AA; 51964 MW; 6D117044 CRC32;

Query Match 12.3% Score 431; DB 4; Length 444;
Best Local Similarity 26.8% Pred. No. 11le-58;
Matches 118; Conservative 112; Mismatches 178; Indels 32; Gaps

Db 13 QGTPRYFTWDEVAQSGCERWLVIDRKVYNISETRRHPGSRVISHYAGODATDPFV 72
QY 2 EGEAKYIYTAEDLRNRKNSGDLWISIQGKYVDCSRWAAEBHPGCVGPLLSLAGQDVTDAFI 61
Db 73 AFHINKGLVKKYWNLSLLIGLSPEQSFSEPTKNKELTDSEFRELRAIVRGLM-KANHYF 131
QY 62 AYH-P-GTAWRHLDPLFTGYVYUKD--FE-V-S-EISKYRRLLENMSRSGIFKEKGHHI 114

```

```

132 Db FLDYLLHILLDGAAMLTTLVWFGTSTFLPCLLCAVLSSAYQAQAGWLQHDFGHLVSFSTSK 191
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
115 QY MTFIVGVAVNMMAIIVGVL-ASESGVGH-MLCGALLGLLWIQAAYVGHDGSHVQVMPTRG 172
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
192 Db WNHLLHHFVIGH-LKAPASWNHMHFQHHAKPNCRKPDDI-NMHPF-FFA-L-CKLIS 246
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
173 QY YNRITO-LIAGNLTIGISIAWKWTNAHHLACNSLDYPDQLQHPVFVAVSTRLENSITS 231
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 Db VELGK-QK-KKMPY--NHQHKYFFLIGPPALLPLFYQWIFYEVIQKKW-DLAWM-I 300
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
232 QY VFGRVLKDEVARFLVSQHHTYYPMIFGRVNLFIQTELL-LLTRRDVPDALNLMI 290
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 Db T-FYVRFFLYTP-LLGLKAFLGLEFIVRFLSNFVVVT-QMNHIP--MHIDHRNDMW 355
    : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
291 QY AVEWTWFPL-FVSCLPWPREFG-FVLISFAVTA-IQHVOFTLNHFSGDTYVGPCKGDNW 347
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
356 Db VSTLOLATCNVHKSASFNDWFSGHNLNQIEHHILFPMPRHNYKHVAPLVOSLCAKHGIEVQ 415
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
348 QY FEKQTGTITDITCPWMDWFFGGQLQEHLHFLPRPGOLRKRIAPLARDLCKKHGMYP 407
    : : : : : : | : | : | : | : | : | : | : | : | : | : | : |
416 Db SKPLLS-AFADIIHSLKEG 434
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
408 QY SGFGWDANVRTLRDA 427
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT      4
ID O61388 PRELIMINARY; PRT; 443 AA.
AC O61388:
DT 01-AUG-1998 ('REMBREL. 07, CREATED)
DT 01-AUG-1998 ('REMBREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 ('REMBREL. 07, LAST ANNOTATION UPDATE)
DE DELTA6-FATTY-ACID-DESATURASE.
CA CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACCELOWATES; NEMATODA; SECERNENTEA; RHABDITIIDA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97268723.
RA SAYANOVA O., SMITH M.A., LAPINSKAS P., STOBART A.K., DOBSON G.,
RA CHRISTIE W.W., SHEURY P.R., NAPIER J.A.:

```


Db 60 EDVGHSDSAREMDKYYIG 78
 YQ 61 IAY-HPGTAWRHLDPLFTG 78

RESULT 9
 ID O04926 PRELIMINARY; PRT; 911 AA.
 AC O04926;
 DT 01-JUL-1997 (TREMBREL. 04, CREATED)
 DT 01-JUL-1997 (TREMBREL. 04, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBREL. 06, LAST ANNOTATION UPDATE)
 DE NADH NITRATE REDUCTASE (EC 1.6.6.3) (NITRATE REDUCTASE (NADPH)).
 GN STNR3.
 OS SOLANUM TUBEROSUM (POTATO).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 CC SOLANALES; SOLANACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DESIREE;
 RA HARRIS N., FOSTER J.M., KUMAR A., DAVIES H.V., WRAY J.L.;
 RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- CATALYTIC ACTIVITY: NADPH + NITRATE = NADP(+) + NITRITE + H(2)O.
 CC -1- COFACTOR: FAD; HEME; MOLYBDENUM.
 DE EMBL; U95317; G1946812;
 DR PROSITE; PS00191; CYTOCHROME_B5; 1.
 DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
 DR PFAM; PF00173; heme_1.
 DR PFAM; PF00174; oxidored_molyb.
 DR PFAM; PF00175; oxidored_fad.
 KW OXIDOREDUCTASE; HEME.
 SQ SEQUENCE 911 AA; 102609 MW; 44AC790F CRC32;

Query Match 5.7%; Score 200; DB 10; Length 911;
 Best Local Similarity 39.7%; Pred. No. 3.04e-15;
 Matches 29; Conservative 16; Mismatches 27; Indels 1; Gaps 1

Db 532 MNTASRWYMSV-VRKHNSDSAWIIVGHYIDASRLKDPHGVDSILLNAGDCTEEF 590
 YQ 1 MEEGAKYITAEIDRLRHNSGDLWISIQGVYDCSWAAEHGPEVPLLSIAGQDVTDAF 60

Db 591 DAIHSDKAKKILE 603
 YQ 61 IAYHPGTAWRHLD 73

RESULT 10
 ID O24390 PRELIMINARY; PRT; 911 AA.
 AC O24390;
 DT 01-JAN-1998 (TREMBREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBREL. 07, LAST ANNOTATION UPDATE)
 DE NADH NITRATE REDUCTASE (EC 1.6.6.3).
 GN STNR2.
 OS SOLANUM TUBEROSUM (POTATO).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 CC SOLANALES; SOLANACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. DESIREE;
 RA HARRIS N., FOSTER J.M., KUMAR A., DAVIES H.V., WRAY J.L.;
 RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; U76701; G1673613;
 DR PROSITE; PS00191; CYTOCHROME_B5; 1.
 DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
 DR PFAM; PF00173; heme_1.
 DR PFAM; PF00174; oxidored_molyb.
 DR PFAM; PF00175; oxidored_fad.
 KW OXIDOREDUCTASE; HEME.
 SQ SEQUENCE 911 AA; 102596 MW; 6AECCE9 CRC32;

Query Match 5.7%; Score 200; DB 10; Length 911;
 Best Local Similarity 39.7%; Pred. No. 3.04e-15;

[illegible]

RESULT	11
ID	Q42342
AC	Q42342;
DT	01-NOV-1996 (TEMBREL.
DT	01-NOV-1996 (TEMBREL.
DT	01-AUG-1998 (TEMBREL.
DE	CITOCROME B5 (FRAGMENT).
OS	ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; CAPPARALES; CRUCIFERAE.

[1] SEQUENCE FROM N. A.
RC TISSUE-CELL SUSPENSION CULTURE OF ECOTYPE COLUMBIA;
RA COOKE R., LAUDIE M., RAYNAL M., DELSENY M.;
RL SUBMITTED (MAR-1996) TO EMBL/GENEBANK/DBJ DATA BANKS.

Query Match 5.5%; Score 194; DB 10; Length 113;
Best Local Similarity 40.5%; Pred. No. 3.17e-14;
Matches 30; Conservative 16; Mismatches 27; Indels

```

5  RKVLSPEEYSKINKTKDCWLLISGKVYDVTFPMFDHDPGGDEVLLSSTGKDANDEEDVCH 64
6  KKYITREDRUNKSGDLWISIQGVYDCSRWAAHEGGEVPLLSAGQDVDTAFAY - H 64

Db      65  STARDMDKIFYG 78
QY      65  PGTAWRHLDPLETG 78

```

RESULT	12
ID	O48845 PRELIMINARY; PRT: 134 AA.
AC	AQ48845;
DT	01-JUN-1998 (TREMREL. 06, CREATED)
DD	01-JUN-1998 (TREMREL. 07, LAST SEQUENCE UPDATE)
DE	01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DZ	POTENTIAL CYTOCHROME B5.
GN	F24L7_14.
OS	ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC	CAPPARALES; CRUCIFERAE.

```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CV, COLUMBIA;
RA      ROUNSEL, S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,
RA      SYKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,
RA      SOMERVILLE C.R., VENTER J.C.;
RL      SUBMITTED (FEB-1998) TO. EMBL/GENBANK/DBJ DATA BANKS.
DR      EMBL; AC003974; G9314701; -.
DR      PROSITE; PS00191; CYTOCHROME_B5; 1.
KW      HEME.
SQ      SEQUENCE 134 AA: 15016 MW: 3EAC1E72 CRC32:

```

Query Match 5.5%; Score 194; DB 10; Length 134;
Best Local Similarity 38.0%; Pred. No. 3.17e-14;
Matches 30; Conservative 16; Mismatches 31; Indels

Db 1 MGDEAKIF-TLSEVSEHNQAHDCWIVINGKVVNTKFLSDHPGGDDVLLSSSTGKDATDDF 59

QY	1	MEGEAKKVI	TAEDLRRHNKSGDLWISIOGKYDCSRWAAEHPPGGEVPLLSLAGQDVTDAF	60
Db	60	EDVGHSESAREMMEQYYVG	78	
QY	61	IAY -HPGTAWRHLDPLFTG	78	
RESULT	13			
ID	O05874	PRELIMINARY;	PRT:	427 AA:

RESULT	13	PRELIMINARY;	PRT;	427 AA.
ID	O05874			
AC	O05874;			
DT	01-JUL-1997	(TREMBREL. 04, CREATED)		
DT	01-JUL-1997	(TREMBREL. 04, LAST SEQUENCE UPDATE)		
DT	01-JUL-1997	(TREMBREL. 04, LAST ANNOTATION UPDATE)		
DE	HYPOTHETICAL 48.4 KD PROTEIN.			
GN	MTCY20B11.04C.			
OS	MYCOBACTERIUM TUBERCULOSIS.			
OC	MYCOBACTERIA; TUBERCUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.			
OC	PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.			

[1] MURRAY, JAMES D., JR. 1990. *Genetics of the*
 SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RA BADCOCK K., CHURCHER C.M.;
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA COLE S.T., BARRELL B.G., RAJANDREAM M.A.;
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBI DATA BANKS

RN [3] SEQUENCE FROM N.A.
 RP STRAIN-H37RV;
 RC MEDLINE; 96181548.
 RX PHILIPP W.J. POULET S. EIGLMEIER K. PASCOPELLA L.
 RA [3]

RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RL PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).

DR EMBL; Z95121; E314467; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 427 AA; 48443 MW; 5D40CABB_CRC32;

Query Match 5.5%; Score 193; DB 2; Length 427;
Best Local Similarity 32.9%; Pred. NO. 4.68e-14;
Matches 26; Conservative 20; Mismatches 32; Indels 1; Gaps 1;

Db 265 FTKTDWIGPKQ-WYLRQMGLGSANFAGPALRFMSGNLCHQIEHHLYDPLPSNRLHEIS 323
:- : ||| :- : | :- : | :- : | :- : | :- : | :- : | :- : | :- : | :- : | :- : | :- :
Qy 333 FSGDTVGPCKDNWFEEKQTGTIDTCTPPMDWFFGCLQFQLEHHLFPRLPRGLQRKIA 392

Db 324 VRVREVCDRYDLPTTGSF 342
| : : | : : | : : |

RESULT	I4
ID	Q43042
AC	PRELIMINARY; PRT; 915 AA.
DT	Q43042;
DT	01-NOV-1996 (TREMBREL_01, CREATED)
DT	01-NOV-1996 (TREMBREL_01, LAST SEQUENCE UPDATE)
DT	01-AUG-1998 (TREMBREL_07, LAST ANNOTATION UPDATE)
DE	NITRATE REDUCTASE APOENZYM (EC 1.6.6.1) (NITRATE REDUCTASE (NADH))
DE	(ASSIMILATORY NITRATE REDUCTASE).
GN	NIA.
OS	PETUNIA HYBRIDA (PETUNIA).
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC	SOLANALES; SOLANACEAE.
FN	[1]

[A]		SEQUENCE FROM N.A.
RP	-	STRAIN-TURLI3 ; TISSUE-LEAF;
RC	-	SALAMABOUT M., BUDANG H.D. ;
RA	-	GENE 0-0-(0).
RL	-	CATALYTIC ACTIVITY: NADH + NITR
CC	-1-	COFACTOR: FAD OR FMN; HEME; MOL
CC	-	ENBL; L11563; G484212; -
DR	-	PROSITE: PS00191; CYTOCHROME P450
DR	-	

W P R F (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 11:33:50 1999; MasPar time 13.14 Seconds
556.518 Million cell updates/sec

ar output not generated.

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pep
Perfect Score: 3515
Sequence: 1 MEGEAKYITAEDLRHNS.....LNSAPCPKILGYGEAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 36.009; Variance 155.761; scale 0.231

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	2134	60.7	448	18	R98455 Borage delta-6-desatu	4.05e-194
2	205	5.8	899	17	R90987 Nitrate reductase.	7.83e-08
3	200	5.7	911	2	R10333 Deduced sequence of t	2.08e-07
4	193	5.5	904	8	R41757 Nitrate reductase Nia	8.15e-07
5	189	5.4	130	23	W22848 Mortierella alpina cy	1.17e-06
6	163	4.6	99	11	R57734 Pre-apo-protein encod	2.52e-04
7	163	4.6	121	11	R57733 Pre-apo-protein encod	2.52e-04
8	163	4.6	156	11	R57732 Bacterial delta-6-des	6.41e-04
9	158	4.5	359	7	R34102 Synchocystis delta-6	1.12e-03
10	135	4.4	439	18	R98456 Fructosyl amino acid	8.96e+00
11	104	3.0	437	25	W22134 IL-4 Stat peptide.	1.24e+01
12	102	2.9	847	15	R83320 Beta-carotene hydroxy	2.79e+01
13	97	2.8	176	3	R13988 CR1 protein	2.79e+01
14	97	2.8	231	1	P92219 Sequence encoded by t	4.49e+01
15	94	2.7	1594	5	P81184 Protein encoded by OR	8.39e+01
16	90	2.6	266	26	W21774 Protein encoded by OR	8.39e+01
17	90	2.6	266	26	W21775 Wild type human epimo	7.19e+01
18	91	2.6	288	13	R66476	

19	91	2.6	288	7	R36549	Human epimorphin.	7.19e+01
20	90	2.6	410	10	R51368	Protein containing Cy	8.39e+01
21	91	2.6	418	11	R60501	Linoleic-acid-desatur	7.19e+01
22	90	2.6	422	15	R88131	Aureobasidin sensitiv	8.39e+01
23	90	2.6	422	15	R88130	Aureobasidin resistan	8.39e+01
24	90	2.6	466	8	R39523	Sequence of gamma-ami	8.39e+01
25	90	2.6	921	5	R28407	CGS-PDE from clone 3C	8.39e+01
26	90	2.6	921	20	W11252	Clone p3CGS-5 cyclic	8.39e+01
27	90	2.6	921	13	R69727	Cyclic-GMP stimulated	8.39e+01
28	90	2.6	921	22	W18048	Cyclic-GMP stimulated	8.39e+01
29	90	2.6	942	13	R69728	Cyclic-GMP stimulated	8.39e+01
30	90	2.6	942	20	W11239	Cyclic GMP stimulated	8.39e+01
31	90	2.6	942	22	W18049	CGS-PDE amino acid se	8.39e+01
32	90	2.6	942	5	R28408	CGS PDE from bovine b	8.39e+01
33	91	2.6	1026	27	W32059	Dogfish shark kidney	7.19e+02
34	88	2.5	109	3	P61525	Sequence encoded by s	1.14e+02
35	88	2.5	114	22	W20571	H. pylori secreted or	1.14e+02
36	88	2.5	125	9	R47656	Interferon induced 1-	1.14e+02
37	88	2.5	352	25	W16646	Truncated HSV specifi	1.14e+02
38	88	2.5	360	25	W16645	Truncated HSV specifi	1.14e+02
39	88	2.5	362	25	W16642	Truncated HSV specifi	1.14e+02
40	88	2.5	375	25	W16644	Truncated HSV specifi	1.14e+02
41	88	2.5	379	25	W16641	Truncated HSV specifi	1.14e+02
42	88	2.5	392	25	W16640	Truncated HSV specifi	1.14e+02
43	88	2.5	412	25	W16639	Truncated HSV specifi	1.14e+02
44	88	2.5	438	25	W16638	Truncated HSV specifi	1.14e+02
45	89	2.5	491	2	P70464	Sequence of gpJ encod	9.80e+01

ALIGNMENTS

RESULT 1
ID R98455 standard; Protein; 448 AA.
AC R98455;
DE 15-SEP-1996 (first entry)
DT Borage delta-6-desaturase.
KW Delta-6-desaturase; gamma-linolenic acid; transgenic plant; borage;
KW polyunsaturated fatty acid; octadecatetraenoic acid;
OS Chilling resistance; oilseed.
OS Borage officinalis.
FH Key Location/Qualifiers
FT region 156..163
FT /label= Lipid_box
FT region 196..200
FT /label= Metal_box-1
FT region 372..377
FT /label= Metal_box-2
FT WO9621022-A2.
PN 11-JUL-1996.
PD 28-DEC-1995; IB1167.
PF 30-DEC-1994; US-366779.
PR (RHON) RHONE POULENC AGROCHIMIE.
PI Freyssinet GL, Nuccio M, Nunberg AN, Reddy AS, Thomas TL;
DR WPI; 96-333937/33.
DR N-PSDB; T30395.
PT Transgenic plants comprising the borage delta-6-desaturase gene
PT show increased production of gamma linolenic acid and having
PT increased resistance to chilling
PS Claim 3; Page 52-53; 73pp; English.
CC Borage delta-6-desaturase (R98455) catalyses the conversion of
CC linoleic acid to gamma-linolenic acid (GLA). Its sequence was
CC deduced from that of the delta-6-desaturase gene (T30395) isolated
CC from a borage membrane-bound polysomal library. The sequence is
CC distinct from that of Synchocystis delta-6-desaturase (R98456).
CC Expression of the desaturase in transgenic plants, esp. sunflower,
CC soybean, maize, tobacco, peanut, carrot or oilseed rape, results in
CC increased GLA prodn. Alteration of the plant membrane lipids as a
CC result of expression of the desaturase may also result in increased
CC resistance to chilling.
SQ Sequence 448 AA;
Query Match 60.7%; Score 2134; DB 18; Length 448;
Best Local Similarity 57.8%; Pred. No. 4.05e-194;

Matches 262; Conservative 105; Mismatches 80; Indels 6; Gaps 6;

Db 1 maaqikkyitsdelknhdpgdglwisiqkaydsvdkdhpqgsfpklsagqevtdaf 60
QY 1 MGEAKKYYITAEEDLRHNSGDLWISIOGVYDCSRWAAEHFPGGEVPLLSLAGQDVTDAF 60

Db 61 vafpbaetwknldkftgyy-lkdyvsvseksdyrklvfefskmglydkkgh-imfatlc 118
QY 61 IAYHPTAWRHLDPLFTGYYIYKDFEVSLSKDYRRLNEMSRSGIFERKKGHHIM-TFV 119

Db 119 fiamfamvvyvlfceegvhlvflfsgclmgflwsgwighdaghymvsvdsrlnkfmgi 178
QY 120 GVAVMAAIVGVGLASVSGVHMLCGALLGLLWIAAAYVGHDSGHVQVMPTRGYNRIITQL 179

Db 179 faancslgisgwknwnhahiaensleydpdlqyfpilvsskffgsltsfhfyeakrit 238
QY 180 IAGNLTGSIAMWKTWTHNAHLACNSLDYDPDLQHIPVFAVSTRFNSITSVYGRVLK 239

Db 239 fgsrlsfvysghwtfypimcaarlmvyslmlltknsvyragellqclvfiwyp 298
QY 240 FDEVARFLVSYQHWTYIPWIFGRVNLFIQTLELLLRDVPDRALNMGIAVFTWTFPL 299

Db 299 lvsclpnwgerimfviaslsvtgmvgqvqfslnhfssvvygkpgnnwfeqtdgtldis 358
QY 300 FVSCLPNWERGFLVLSFAVTAIOHVQTLNHFSGDTYVGPCKDNWFEKTKGTIDIT 359

Db 359 cppwmdfhgglgfiehllfpmprcnrlkispvyielckhnlpy-nyasfskanent 417
QY 360 CPPWMDWFEGGLQFLEHFLFPLRGQLRKIAPLARDLCKKHGMYRSGFGEWDDANVRT 419

Db 418 lrltntalqarditk-plpknlw-ealhtg 448
QY 420 IRTLDAVAQVARDLNSAPCKPLGLVGEAYNTHG 452

RESULT 2

ID R90987 standard; Protein; 899 AA.
AC R90987;
DT 01-AUG-1996 (first entry)
DE Nitrate reductase.
KW Nitrate reductase; populus nigra; absorption; nitrogen oxide; pollutant.
OS Populus nigra.
PN J08023978-A.
PD 30-JAN-1996.
PF 14-JUL-1994; JP-162197.
PR 14-JUL-1994; JP-162197.
PA (TOY) TOYOTA CHUO KENYUSHO KK.
DB (TOYT) TOYOTA JTDOSHA KK.
WP: 96-133419/14.
DR N-PSDB; T12999.
PT Populus sp. nitrate reductase gene - useful for generating
PT transgenic plants with enhanced nitrogen oxide absorbing activity
PS Claim 1; Page 4-8; 9pp; Japanese.
CC This sequence represents the nitrate reductase sequence obtained from
CC Populus nigra L. The gene encoding this sequence can be used to generate
CC plants, especially trees, having enhanced activity for absorbing nitrogen
CC oxide pollutants.
SQ Sequence 899 AA.

Query Match 5.8%; Score 205; DB 17; Length 899;
Best Local Similarity 38.4%; Pred. No. 7.83e-08;
Matches 28; Conservative 17; Mismatches 27; Indels 1; Gaps 1;

Db 527 mntssktfsmag-vkksadsawilvhghvdydctfrikdhpqgtdslinagtacteeef 585
QY 1 MGEAKKYYITAEEDLRHNSGDLWISIOGVYDCSRWAAEHFPGGEVPLLSLAGQDVTDAF 60

Db 586 daihsdkakmlie 598
QY 61 IAYHPTAWRHLD 73

RESULT 3

ID R10333 standard; Protein; 911 AA.
AC R10333;
DT 08-APR-1991 (first entry)
DE Deduced sequence of tomato nitrate reductase.
KW tomato nitrate reductase; nitrogen assimilation.
OS Lycopersicon esculentum.
PN EP-409730-A.
PD 23-JAN-1991.
PF 18-JUL-1990; 402077.
PR 19-JUL-1989; FR-009707.
PA (INRG) INST NAT RECH AGRON.
PI Daniel-Vedele F, Caboche M;
WPI: 91-024287/04.
DR Q-PSDB; Q10280.
PT New DNA encoding tomato nitrate reductase - and related cloning
PT and expression vectors, used to improve nitrogen assimilation in
PT plants
PS Claim 2; Fig 1; 27pp; French.
CC An EcoRI digest of tomato DNA was screened with labelled tobacco
CC nitrate reductase cDNA under low stringency conditions. A 6.5Kb
CC fragment contained part of the tomato nitrate reductase gene and
CC was used, under high stringency, to screen a Hind III library. A
CC 7kb fragment was isolated, having 1.5kb in common with the EcoRI
CC fragment. Restriction fragments were subcloned and sequenced. The
CC deduced protein sequence is encoded by four exons. Comparison with
CC known nitrate reductases from other plants, e.g. Nicotiana,
CC Solanum, Trifolium pratense, Cucumis sativus, etc., shows conserved
CC regions. These have structural and/or functional importance and
CC tend to be separated by variable regions, often hydrophilic, which
CC are possibly specific surface epitopes of the enzyme.
SQ Sequence 911 AA;

Query Match 5.7%; Score 200; DB 2; Length 911;
Best Local Similarity 39.7%; Pred. No. 2.08e-07;
Matches 29; Conservative 16; Mismatches 27; Indels 1; Gaps 1;

Db 532 mntaskmymse-vrkhnssdsawilvhghydasflkdhpqgvdslinagtcteeef 590
QY 1 MGEAKKYYITAEEDLRHNSGDLWISIOGVYDCSRWAAEHFPGGEVPLLSLAGQDVTDAF 60

Db 591 daihsdkakmlie 603
QY 61 IAYHPTAWRHLD 73

RESULT 4

ID R41757 standard; Protein; 904 AA.
AC R41757;
DT 21-MAR-1994 (first entry)
DE Nitrate reductase Nia2 derivative.
KW Nitrate reductase; germination; flowering; ripening; development;
KW growth stimulation; Agrobacterium tumefaciens; nitrate.
OS Nicotiana sp.
PN W09318154-A.
PD 16-SEP-1993.
PF 05-MAR-1993; F00222.
PR 05-MAR-1992; FR-002658.
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Caboche M, Chupau Y, Dorlhac F, Morot-gaudry J;
PI Vincentz M;
DR WPI: 93-303468/38.
DR P-PSDB; Q48468.
PT Inducing over-expression of nitrate reductase in plants - esp. by
PT incorporation of foreign gene, for stimulating early development
PT and reducing nitrate accumulation
PS Disclosure; Figure 3; 37pp; English.
CC Overexpression of nitrate reductase (NR) can stimulate the early
CC development of plants, shortening the duration of the vegetative
CC phase and causing earlier germination, flowering and ripening by
CC about two weeks. Overexpression of NR can also cause the level of
CC nitrate stored in a plant to be reduced, reducing risks to health
CC and also possibly improving organoleptic qualities. The NR gene was
CC introduced into plants by transforming a strain of Agrobacterium

CC tumefaciens with a plasmid and using it to infect the subject
 CC plant(s).
 SQ Sequence 904 AA;

Query Match 5.5%; Score 193; DB 8; Length 904;
 Best Local Similarity 35.6%; Pred. No. 8.15e-07;
 Matches 26; Conservative 19; Mismatches 27; Indels 1; Gaps 1;

Db 527 mntaskmysnse-vrkhsadsawilvhghydatrfkdhpggtdsillnagtdcteeef 585
 QY 1 MEGEAKKIITAEELRRHNSGDLWISIQGKVDCSRWAAEHPGVEPLLSLAGQDVTDFA 60

Db 586 dathsdakkille 598

QY 61 IAYHPTAWRHLD 73

RESULT 5

ID W22848 standard; Protein; 130 AA.

AC W22848;

DT 07-OCT-1997 (first entry)

K Mortierella alpina cytochrome b5.

KW cytochrome b5; cytb5; preparation; production; human;

OS fatty acid.

OS Mortierella alpina.

PN J09121873-A.

PD 13-MAY-1997.

PF 01-AUG-1996; 203735.

PR 01-AUG-1995; JP-196868.

PA (SUNR) SUNTORY LTD.

DR WPI: 97-314231/29.

DR N-PSDB; T75438.

PT Mortierella alpina cytochrome b5 (cytb5) gene - used for recombinant

PT production of cytb5 which is used in the preparation of essential

PT fatty acids

PS Claim 3; Page 11; 14pp; Japanese.

CC The present sequence is the Mortierella alpina cytochrome b5

CC (cytb5), useful for the efficient preparation of human essential

CC fatty acids.

SQ Sequence 130 AA;

Query Match 5.4%; Score 189; DB 23; Length 130;

Best Local Similarity 38.0%; Pred. No. 1.77e-06;

Matches 30; Conservative 16; Mismatches 31; Indels, 2; Gaps 2;

Db 2 aelksf-tladsghtkdsilylaihkvdyctgfidehpggeevlidesgrdatesfed 60

QY 3 GEAKKIITAEELRRHNSGDLWISIQGKVDCSRWAAEHPGVEPLLSLAGQDVTDFA 62

QY 61 vghsdeardimskllvgef 79

QY 63 Y-HPGTAWRHLDPLFTGY 80

RESULT 6

ID R57734 standard; Protein; 99 AA.

AC R57734;

DT 24-MAR-1995 (first entry)

DE Pre-apo-protein encoded by plasmid pSEC-cyt/N

KW Plasmid pSEC-cyt/N; vector; holo-protein; pre-apo-protein;

KW protein synthesis; Escherichia coli; cytoplasm cytochrome.

OS Synthetic.

OS Mortierella alpina.

PN W09417191-A.

PD 04-AUG-1994.

PF 27-JAN-1994; G00161.

PR 27-JAN-1993; GB-001553.

PA (UYWA-) UNIV COLLEGE WALES.

PI Kaderbhai MA;

DR WPI: 94-264104/32.

DR N-PSDB; Q67222.

PT Genetic precursor unit causing periplasmic translocation of

PT pre-apo-protein - for processing the halo-protein, useful in

PT protein synthesis, assay of signal peptidase and identification

PT of its inhibitors

PS Claim 10; Fig.6; 32pp; English.

CC This sequence encodes the pre-form of an apo-protein. The pre-apo-

CC protein is translocated from a cytoplasmic cell region to a

CC periplasmic region where formation of process apo-protein and

CC conversion to halo-protein occur. The protein is expressed in the

CC cytoplasm of E. coli, especially TB-1 and N4830-1. The apo-protein

CC is a cytoplasmic cytochrome, especially having a soluble core domain

CC of cytochrome-b5 of liver endoplasmic reticulum.

SQ Sequence 99 AA;

Query Match 4.6%; Score 163; DB 11; Length 99;

Best Local Similarity 37.1%; Pred. No. 2.52e-04;

Matches 26; Conservative 13; Mismatches 31; Indels 0; Gaps 0;

Db 10 kyytleeqhkhdkstwtvllhkhvdytlkfleehepggeevlreqagagdatenfedvghs 69

QY 7 KYITAEELRRHNSGDLWISIQGKVDCSRWAAEHPGVEPLLSLAGQDVTDFAIYHPG 66

Db 70 tdarelskty 79

QY 67 TAWRHLDPLF 76

RESULT 7

ID R57732 standard; Protein; 121 AA.

AC R57732;

DT 24-MAR-1995 (first entry)

DE Pre-apo-protein encoded by plasmid pAA-cyt.

KW Plasmid pAA-cyt; vector; holo-protein; pre-apo-protein;

KW protein synthesis; Escherichia coli; cytoplasm cytochrome.

OS Synthetic.

FH Key

FT peptide

FT region

FT /note= "alkaline phosphatase signal peptide"

FT /note= "soluble core of cytochrome-b5"

PN W09417191-A.

PD 04-AUG-1994.

PF 27-JAN-1994; G00161.

PR 27-JAN-1993; GB-001553.

PA (UYWA-) UNIV COLLEGE WALES.

PI Kaderbhai MA;

DR WPI: 94-264104/32.

DR N-PSDB; Q67220.

PT Genetic precursor unit causing periplasmic translocation of

PT pre-apo-protein - for processing the halo-protein, useful in

PT protein synthesis, assay of signal peptidase and identification

PT of its inhibitors

PS Claim 10; Fig.4; 32pp; English.

CC This sequence encodes the pre-form of an apo-protein. The pre-apo-

CC protein is translocated from a cytoplasmic cell region to a

CC periplasmic region where formation of process apo-protein and

CC conversion to halo-protein occur. The protein is expressed in the

CC cytoplasm of E. coli, especially TB-1 and N4830-1. The apo-protein

CC is a cytoplasmic cytochrome, especially having a soluble core domain

CC of cytochrome-b5 of liver endoplasmic reticulum.

SQ Sequence 121 AA;

Query Match 4.6%; Score 163; DB 11; Length 121;

Best Local Similarity 37.1%; Pred. No. 2.52e-04;

Matches 26; Conservative 13; Mismatches 31; Indels 0; Gaps 0;

Db 32 kyytleeqhkhdkstwtvllhkhvdytlkfleehepggeevlreqagagdatenfedvghs 91

QY 7 KYITAEELRRHNSGDLWISIQGKVDCSRWAAEHPGVEPLLSLAGQDVTDFAIYHPG 66

Db 92 tdarelskty 101

QY 67 TAWRHLDPLF 76

RESULT 8

```

Best Local Similarity 31.68; Pred.No. 6.41e-04; Gaps 3;
Matches 25; Conservative 16; Mismatches 35; Indels 3; Gaps 3;

Db      273 dawaioqirrtanfatnmpfw-nwfegglnhgvtlhlfpnichihyppqlenikdvcegf 331
       |:| : | : : | | : | : ||||| : : : : | : | : | : | : | : | : |
Qy      345 DNWFERKTKGTIDI-TCCP-WMDWFFGGQLFQLEHLFLPRLPQGRLRKIAPLARDLCCKH 402

Db      332 gveykvyptfkaaiastyr 350
       |:| : | : | | | | |
Qy      403 GMPYRSFGFWDDANVRIR 421

RESULT 10
ID     R98456 standard; Protein; 359 AA.
AC     R98456;
DT     D
DE     Synecocystis delta-6-desaturase.
KW     Delta-6-desaturase; gamma-linolenic acid; transgenic plant; borage;
KW     polyunsaturated fatty acid; octadecatetraenoic acid;
KW     chilling resistance; oilseed.
KW     Synecocystis sp. strain PCC 6803 (ATCC 27184).
FH     Key Location/Qualifiers
FT     region 85..92 ..
FT     /label= Lipid_box
FT     region 123..128
FT     /label= Metal_box-1
FT     region 302..307
FT     /label= Metal_box-2
PN     WO9621022-A2.
PD     11-JUL-1996.
PF     28-DEC-1995; IB1167.
PP     30-DEC-1994; US-366779.
PR     (RHON ) RHONE POULENC AGROCHIMIE.
PI     Freyssinet GL, Nuccio M, Nunberg AN, Reddy AS, Thomas TL;
DR     WPI; 96-333997/33.
DR     N-PSDB; T30396.
PT     Transgenic plants comprising the borage delta-6-desaturase gene -
PT     show increased production of gamma linolenic acid and having
PT     increased resistance to chilling
PS     Disclosure; Page 48-49; 75pp; English.
SC     The conversion of linoleic acid to gamma-linolenic acid (GLA). Its
CC     sequence was deduced from that of the delta-6-desaturase gene
CC     (T30396) isolated from a Synecocystis cosmid genomic library.
CC     The sequence is distinct from that of borage delta-6-desaturase
CC     (R98456). Expression of the desaturase in bacteria, fungi,
CC     animals or plants results in increased GLA prodn. Alteration of
CC     the transgenic plant membrane lipids as a result of expression of
CC     the desaturase may result in increased resistance to chilling.
SQ     Sequence 359 AA;

Query Match 4.4%; Score 155; DB 18; Length 359;
Best Local Similarity 32.88; Pred.No. 1.12e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db      273 dawaioqirrtanfatnmpfw-nwfegglnhgvtlhlfpnichihyppqlenikdvcegf 331
       |:| : | : : | | : | : ||||| : : : : | : | : | : | : | : | : |
Qy      345 DNWFERKTKGTIDI-TCCP-WMDWFFGGQLFQLEHLFLPRLPQGRLRKIAPLARDLCCKH 402

Db      332 gveykyv 338
       |:| : | : | | | | |
Qy      403 GMPYRSF 409

RESULT 11
ID     W24134 standard; Protein; 437 AA.
AC     W24134;
DT     DT
DE     Fructosyl amino acid oxidase.
DE     Fructoseyl amino acid oxidase.
KW     Fructoseyl amino acid oxidase; FAOD-P; enzyme; amadori compound;
KW     alpha-ketoaldehyde amine derivative; amadori detection.
OS     Penicillium janthinellum.
PN     WO9721818-A1.
```

```

Query Match          2.9%; Score 102; DB 15; Length 847;
Best Local Similarity 21.1%; Pred. No. 1.24e+01;
Matches 24; Conservative 36; Mismatches 44; Indels 10; Gaps 5

Db 433 sendryp-fvvaarvpwkmcetlnlkfnaevtnrgllp-ehflflaqlkfndnslsme 490
      ::::| | | | | : : : : | | | | | : : : | | : | | : | : |
QY 211 PDLQHIFVFAVSTRL-FNSITSFYGRVLKDFEVARFLVSQHWHTYVPVMIFGRVNLFIQ 269
      : | : : : : : : : : : | | | | | : : : | | : | | : | : |
Db 491 afchrsvswsfknel-llgrgtfgqwdgvdldtkrcrlsrwysdarlligfls 543
      : | : : : : : : : : : | | | | | : : : | | : | | : | : |
QY 270 TFLLLTRRDVPDRALNMGIA-VFWTFP--L-FVS-CLPN-WPERGFEVLIS 317
      : | : : : : : : : : : | | | | | : : : | | : | | : | : |

RESULT 13
ID R13988 standard; Protein; 176 AA.
AC R13988;
DT 26-NOV-1991 (first entry)
DE Beta-carotene hydroxylase - variant.
KW GGPP; carotenoid; phytoene; zeaxanthin; lycopene; ss.
OS Erwinia herbicola EHO-10 (E. vulneris - ATCC 39368).
PN W09113078-A.
PD 05-SEP-1991.
PF 04-AUG-1991; U01458.
PR 02-MAR-1990; W0-487613.
PR 18-MAY-1990; U5-52551.
PR 03-AUG-1990; U5-562674.
PA (STAD.) AMOCO CORP.
PI Ausich RL, Brinkhaus FL, Mukharji I, Proffitt JH, Yarger JG;
PI Yen HC;
DR WPI; 91-281410/38.
DR N-PSDB; Q13725.
PT Biosynthesis of carotenoid(s) in genetically engineered hosts -
PT using DNA encoding enzymes from Erwinia herbicola
PS Disclosure; Fig 2(1-3); 313pp; English.
CC There are a total of six relevant genes in a 7900 bp region that
CC cause E. coli cells to produce GGPP and the carotenoids phytoene
CC through zeaxanthin diglucoside, which is the final prod. identified
CC in the carotenoid pathway contd. in plasmid pARC376 (contg. a ca.
CC 13 kb chromosomal DNA fragment isolated by Perry et al., J. Bacteriol.,
CC 168:507 (1986)). The genes for geranylgeranyl pyrophosphate (GGPP)
CC synthase, phytoene synthase, phytoene dehydrogenase-4H, lycopene
CC cyclase, beta-carotene hydroxylase, and zeaxanthin glycosylase are
CC represented in Q13716, Q13718, Q13719, Q13722, Q13724 and Q13726
CC respectively.
CC The native sequence (Q13724) was genetically engineered.
CC At the 5' end of the gene, codons encoding the native second and
CC third amino acid have been changed from Leu-val to Val-Leu.
CC Recombinant expression plasmids can be used to produce large amts.
CC of the enzymes and hence large amts. of the carotenoids which they
CC synthesize.
SQ Sequence 176 AA;

Query Match          2.8%; Score 97; DB 3; Length 176;
Best Local Similarity 28.8%; Pred. No. 2.79e+01;
Matches 23; Conservative 22; Mismatches 30; Indels 5; Gaps 5;

Db 38 hhtprkrgvfelndlfav-fvagvaialiavgtagvplqwigcgmtvgyllfl-vhdgl 95
      : : | | | | : : | | | | : : | | | | : : | | | | : : | | : :
QY 99 NEMSRSGIFPKKGHHIMTIFVGVAWMAAI-VYGVLAESVGVHMLCGALLGLWTQAAAY 157
      : : : : : : : : : : | | | | : : | | : : | | : | : : | :

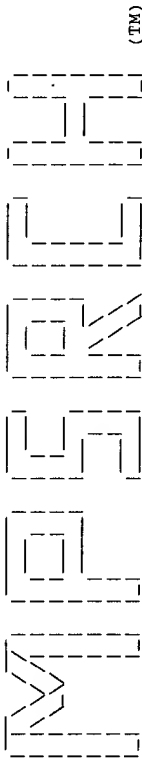
Db 96 v-hgrwpfhwirpylkrll 114
      | | : : | | | | :
QY 158 VGHDSGHYQWMPTRGY-NRI 176
      | | : : | | | | :

RESULT 14
ID P92219 standard; protein; 2317 AA.
AC P92219;
DT 22-FEB-1990 (first entry)
DE CR1 protein
DE Complement; cofactor.
KW Homo sapiens (human).
FH Key
      Location/Qualifiers

```

	Query Match	2.7%	Score 94;	DB 5;	Length 1594;
	Best Local Similarity	26.4%;	Pred. NO. 4.49e+01;		
	Matches 24;	Conservative	24;	Mismatches 36;	Indels 7; Gaps 7;
Db	817 vl-ldeltrylvayit-hhyqvixalkmlvmsfilxrhvmxahkrll-lm-vpxiel	872			
Qy	: :: : :	:	:	: :	: : :::
	237 VLKFEDEVA-RFLSYQHWTYYPMVIGFVNLFQTPELLLTTRDVPDRALNMGIAVFMT	295			
Db	873 xlpbtvncvxvhigqrhlfttlylitqvr	903			
	:: : : :	:	:	: :	: : :::
Qy	296 WFFLFWSC-LPNWPFR-FGFVLISFAVTAIQ	324			

PS inflammation, myelodysplastic disorders, etc.
 CC Claim 11: fig. 1; 191pp; English.
 CC This is full-length CR1 protein. The proteins and fragments bind C3b and/or
 CC the transmembrane region. The proteins and fragments bind C3b and/or
 CC C4b, have cofactor activity and inhibit C3 and C5 convertase activity.
 CC In the sequence, x=untranslated region. This has 7 short consensus
 CC repeats (CR) of 4 long direct homologous repeats (LHR) (see fig. 10 in
 CC patent). Deletion mutants, eg LHR-A for binding of C4b, and LHR-C
 CC for C3b binding, be made, and LHR-B and -C can detect CR2 sequences.
 CC They are useful in diagnosing and treating immune disorders, and prevent
 CC perfusion injury.
 SQ Sequence 2317 AA;



 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 11:38:16 1999; MasPar time 6.00 Seconds

Par output not generated. 712.828 Million cell updates/sec

Title: >US-08-934-254-27
 Description: (1-452) from US08934254.pep
 Perfect Score: 3515
 Sequence: 1 MEGEAKYITAEDLRHNS.....LNSAPCPKILGYGEAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 100342 seqs, 9469514 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfiles1

Statistics: Mean 33.831; Variance 160.640; scale 0.211

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2121	60.3	448	1	US-08-366- Sequence 5, Applicatio	1.93e-176
2	2121	60.3	448	2	US-08-789- Sequence 5, Applicatio	1.93e-176
3	187	5.3	155	2	US-08-801- Sequence 1, Applicatio	6.01e-06
4	171	4.9	134	2	US-08-801- Sequence 4, Applicatio	1.03e-04
5	170	4.8	142	2	US-08-801- Sequence 3, Applicatio	1.22e-04
6	155	4.4	359	1	US-08-366- Sequence 2, Applicatio	1.67e-03
7	155	4.4	359	1	US-08-789- Sequence 2, Applicatio	1.67e-03
8	155	4.4	359	1	US-08-307- Sequence 2, Applicatio	1.67e-03
9	155	4.4	359	1	US-08-478- Sequence 2, Applicatio	1.67e-03
10	155	4.4	359	1	US-08-473- Sequence 2, Applicatio	1.67e-03
11	137	3.9	104	2	US-08-801- Sequence 5, Applicatio	3.57e-02
12	102	2.9	847	1	US-08-781- Sequence 2, Applicatio	1.01e+01
13	102	2.9	847	1	US-08-276- Sequence 2, Applicatio	1.01e+01
14	97	2.8	175	1	US-08-624- Sequence 6, Applicatio	2.17e+01
15	97	2.8	176	1	US-08-096- Sequence 18, Applicati	2.17e+01
16	91	2.6	288	2	US-08-628- Sequence 3, Applicatio	5.33e+01
17	91	2.6	288	1	US-08-690- Sequence 3, Applicatio	5.33e+01
18	91	2.6	418	3	PCT-US94-0 Sequence 72, Applicatio	5.33e+01
19	93	2.6	450	2	US-08-818- Sequence 6, Applicatio	3.96e+01
20	91	2.6	473	1	US-08-439- Sequence 4, Applicatio	5.33e+01
21	91	2.6	473	1	US-08-440- Sequence 3, Applicatio	5.33e+01
22	93	2.6	553	1	US-08-565- Sequence 6, Applicatio	3.96e+01
23	90	2.6	921	3	PCT-US92-0 Sequence 39, Applicati	6.17e+01

24	90	2.6	921	2	US-08-455- Sequence 39, Applicati	6.17e+01
25	90	2.6	921	1	US-08-297- Sequence 39, Applicati	6.17e+01
26	90	2.6	921	2	US-08-455- Sequence 39, Applicati	6.17e+01
27	90	2.6	921	2	US-08-479- Sequence 39, Applicati	6.17e+01
28	90	2.6	921	1	US-07-872- Sequence 39, Applicati	6.17e+01
29	90	2.6	921	1	US-08-297- Sequence 39, Applicati	6.17e+01
30	90	2.6	942	1	US-07-872- Sequence 43, Applicati	6.17e+01
31	90	2.6	942	1	US-08-297- Sequence 43, Applicati	6.17e+01
32	90	2.6	942	3	PCT-US92-0 Sequence 43, Applicati	6.17e+01
33	90	2.6	942	2	US-08-455- Sequence 43, Applicati	6.17e+01
34	90	2.6	942	1	US-08-297- Sequence 43, Applicati	6.17e+01
35	90	2.6	942	2	US-08-479- Sequence 43, Applicati	6.17e+01
36	90	2.6	942	2	US-08-455- Sequence 43, Applicati	6.17e+01
37	88	2.5	125	3	PCT-US93-0 Sequence 17, Applicati	8.27e+01
38	88	2.5	266	2	US-08-773- Sequence 4, Applicatio	8.27e+01
39	89	2.5	383	3	PCT-US94-0 Sequence 2, Applicatio	7.15e+01
40	89	2.5	516	1	US-08-356- Sequence 4, Applicatio	7.15e+01
41	87	2.5	1078	1	US-08-485- Sequence 7, Applicatio	9.57e+01
42	87	2.5	1078	2	US-08-484- Sequence 7, Applicatio	9.57e+01
43	87	2.5	1078	2	US-08-480- Sequence 7, Applicatio	9.57e+01
44	87	2.5	1085	2	US-08-480- Sequence 5, Applicatio	9.57e+01
45	87	2.5	1088	1	US-08-485- Sequence 6, Applicatio	9.57e+01

ALIGNMENTS

RESULT 1
 ID US-08-366-779-5 STANDARD; PRT; 448 AA.
 XX
 AC xxxxxx
 XX
 DT
 DE
 DE Sequence 5, Application US/08366779
 XX
 XX Sequence 5, Application US/08366779
 CC
 CC Patent No. 5614393
 CC GENERAL INFORMATION:
 CC APPLICANT: Thomas, Terry L.
 CC APPLICANT: Reddy, Avutu S.
 CC APPLICANT: Nuccio, Michael
 CC APPLICANT: Freysinet, Georges L.
 CC APPLICANT: Nunberg, Andrew N.
 CC TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A
 CC TITLE OF INVENTION: DELTA 6-DESATURASE
 CC NUMBER OF SEQUENCES: 25
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Scully, Scott, Murphy & Presser
 CC STREET: 400 Garden City Plaza
 CC CITY: Garden City
 CC STATE: New York
 CC COUNTRY: United States
 CC ZIP: 11530
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/366,779
 CC FILING DATE: 30-DEC-1994
 CC CLASSIFICATION: 800
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Presser, Leopold
 CC REGISTRATION NUMBER: 19,827
 CC REFERENCE/DOCKET NUMBER: 83832YXW
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (516) 742-4343
 CC TELEFAX: (516) 742-4366
 CC TELEX: 230 901 SANS UR
 CC INFORMATION FOR SEQ ID NO: 5:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 448 amino acids

```

CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      SEQUENCE 448 AA: 51599 MW: 1110962 CN;

Query Match
Best Local Similarity 60.3%; Score 2121; DB 1; Length 448;
Matches 261; Conservative 105; Mismatches 81; Indels 6; Gaps 6;

Db 1 MAAQIKKYYTSDLNKNDKPGDLWISIOGKAYDVSDWVDPGSGPPLKSLAGQEVTDFAF 60
QY 1 MGEAKKYYTADLRHNSGDLWISIOGKAYDVSDWVDPGSGPPLKSLAGQEVTDFAF 60

Db 61 VAFHPASTWKNLDFKFTGY-LKDYSSVSESKDYKRLVFEFSKMGLYDKKGH-IMPATIC 118
QY 61 IAYHPGTAWRHLDPLFTGYLKKDFEVSISKDYRLLNEMSRSGIFKKGHIMW-TFV 119

Db 119 FIAMLFAMSVYGLVFCGVLHFLSGCLMGFLWISQGWIGHAGHYMVYSDSLNKFEMI 178
QY 120 GVAVMAAIYGVGLASESVGVHMLCGLLGLLWIAQAAVYGHDSGHYQVMPTRGYNRI 179

Db 179 FAANCLSGISIGWKNWNAHIAACNSLEYDPDLOVYIPFLVSSKFFGSLTSHFEYKRLT 238
QY 180 IAGNLTGISAIAWKWTHNAHLACNSLDYDPDLOHPIVFAVSTRFNSITSVFYGRVLK 239

Db 239 FDSLRSFFVSYOHTWTFYIMCAARLNMYVQSLIMLTTRNVSRAQELLCGLVFSIWP 298
QY 240 FDEVARFLSYOHTWTFYPMIFGRVNLFTQTFLTLTRDVPDRALNMGIAVFWTWEPL 299

Db 299 LVSCLPNNGERIMFVIALSVTGMQOVQFSLNHFSSSVTVGPKGNWPEKOTDGTLDIS 358
QY 300 FVSCLPNWPREFGVLIISFAVTAIQHVQFTLNHFSGDTYVGPVPGKGNWPEKOTKGTIDIT 359

Db 359 CPPWMDWFFGGSGFOIEHLLFPKMPRCNLRKISPYVIELCKHNLPLY-NYASFSKANEMT 417
QY 360 CPPWMDWFFGGSGFOIEHLLFPKMPRCNLRKISPYVIELCKHNLPLY-NYASFSKANEMT 419

Db 418 LRTLNTALQARDITK-PLPKNLVW-EALHTHG 448
QY 420 IRTLDAVAQARDLNSAPCPKKLGYGEAYNTHG 452

RESULT 3
ID US-08-934-254-27.1 STANDARD; PRT; 155 AA.
XX
AC xxxxxx
XX
DT
XX
DE
XX
CC Sequence 1, Application US/08801972
CC Sequence 1, Application US/08801972
CC Patent No. 5831018
CC GENERAL INFORMATION:

Sequence 5, Application US/08789936
Sequence 5, Application US/08789936
Patent No. 5789220
GENERAL INFORMATION:
APPLICANT: Thomas, Terry L.
APPLICANT: Reddy, Avutu S.
APPLICANT: Nuccio, Michael
APPLICANT: Freysinet, Georges L.
APPLICANT: Nunberg, Andrew N.
TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A
TITLE OF INVENTION: DELTA 6-DESATURASE
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/789,936
CC      FILING DATE: 28-JAN-1997
CC      CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: 08/366,779
CC      FILING DATE: 30-DEC-1994
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Presser, Leopold
CC      REGISTRATION NUMBER: 19,827
CC      REFERENCE/DOCKET NUMBER: 83832YXW
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (516) 742-4343
CC      TELEFAX: (516) 742-4366
CC      TELEX: 230 901 SANS UR
CC      INFORMATION FOR SEQ ID NO: 5:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 448 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE:
SQ      SEQUENCE 448 AA: 51599 MW: 1110962 CN;

Query Match
Best Local Similarity 60.3%; Score 2121; DB 2; Length 448;
Matches 261; Conservative 105; Mismatches 81; Indels 6; Gaps 6;

Db 1 MAAQIKKYYTSDLNKNDKPGDLWISIOGKAYDVSDWVDPGSGPPLKSLAGQEVTDFAF 60
QY 1 MGEAKKYYTADLRHNSGDLWISIOGKAYDVSDWVDPGSGPPLKSLAGQEVTDFAF 60

Db 61 VAFHPASTWKNLDFKFTGY-LKDYSSVSESKDYKRLVFEFSKMGLYDKKGH-IMPATIC 118
QY 61 IAYHPGTAWRHLDPLFTGYLKKDFEVSISKDYRLLNEMSRSGIFKKGHIMW-TFV 119

Db 119 FIAMLFAMSVYGLVFCGVLHFLSGCLMGFLWISQGWIGHAGHYMVYSDSLNKFEMI 178
QY 120 GVAVMAAIYGVGLASESVGVHMLCGLLGLLWIAQAAVYGHDSGHYQVMPTRGYNRI 179

Db 179 FAANCLSGISIGWKNWNAHIAACNSLEYDPDLOVYIPFLVSSKFFGSLTSHFEYKRLT 238
QY 180 IAGNLTGISAIAWKWTHNAHLACNSLDYDPDLOHPIVFAVSTRFNSITSVFYGRVLK 239

Db 239 FDSLRSFFVSYOHTWTFYIMCAARLNMYVQSLIMLTTRNVSRAQELLCGLVFSIWP 298
QY 240 FDEVARFLSYOHTWTFYPMIFGRVNLFTQTFLTLTRDVPDRALNMGIAVFWTWEPL 299

Db 299 LVSCLPNNGERIMFVIALSVTGMQOVQFSLNHFSSSVTVGPKGNWPEKOTDGTLDIS 358
QY 300 FVSCLPNWPREFGVLIISFAVTAIQHVQFTLNHFSGDTYVGPVPGKGNWPEKOTKGTIDIT 359

Db 359 CPPWMDWFFGGSGFOIEHLLFPKMPRCNLRKISPYVIELCKHNLPLY-NYASFSKANEMT 417
QY 360 CPPWMDWFFGGSGFOIEHLLFPKMPRCNLRKISPYVIELCKHNLPLY-NYASFSKANEMT 419

Db 418 LRTLNTALQARDITK-PLPKNLVW-EALHTHG 448
QY 420 IRTLDAVAQARDLNSAPCPKKLGYGEAYNTHG 452

RESULT 3
ID US-08-934-254-27.1 STANDARD; PRT; 155 AA.
XX
AC xxxxxx
XX
DT
XX
DE
XX
CC Sequence 1, Application US/08801972
CC Sequence 1, Application US/08801972
CC Patent No. 5831018
CC GENERAL INFORMATION:

```

[illegible]


```
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Presser, Leopold  
CC REGISTRATION NUMBER: 19,827  
CC REFERENCE/DOCKET NUMBER: 83832YXW  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (516) 742-4343  
CC TELEXFAX: (516) 742-4366  
CC TELETYPE: 230 901 SANS UR  
CC TYPE: amino acids  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 359 AA; 41425 MW; 698417 CN;  
  
Query Match      4.4%; Score 155; DB 1: Length 359;  
Best Local Similarity 32.8%; Pred. No. 1.67e+03;  
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps  
  
Db 273 DEWAICQIRITANPATNNPFW-NWFCGGLNQHTVTHLFPNCHIHYPQLDENIKDVCQE 331  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 345 DNWEKTKGTGIDI-TCP-PWMDFEGLGFQLEHHLFLPLPGQLRKIAPLDLCCKKH 402  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
  
Db 332 GVEYKVY 338  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 403 GMPYRSF 409  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
  
RESULT          7  
ID US-08-789-936-2 STANDARD; PRT: 359 AA.  
XX xxxxxx  
AC  
AD  
DT  
DT  
XX  
  
DE Sequence 2, Application US/08789936  
XX  
Sequence 2, Application US/08789936  
Patent No. 5789220  
GENERAL INFORMATION:  
CC APPLICANT: Thomas, Terry L.  
CC APPLICANT: Reddy, Avutu S.  
CC APPLICANT: Nuccio, Michael  
CC APPLICANT: Freyssonet, Georges L.  
CC APPLICANT: Nunberg, Andrew N.  
CC TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A  
CC TITLE OF INVENTION: DELTA 6-DESATURASE  
CC NUMBER OF SEQUENCES: 25  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Scully, Scott, Murphy & Presser  
CC STREET: 400 Garden City Plaza  
CC CITY: Garden City  
CC STATE: New York  
CC COUNTRY: United States  
CC ZIP: 11530  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/789,936  
CC FILING DATE: 28-JAN-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/366,779  
CC FILING DATE: 30-DEC-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Presser, Leopold  
CC REGISTRATION NUMBER: 19,827  
CC REFERENCE/DOCKET NUMBER: 83832YXW  
CC TELECOMMUNICATION INFORMATION:
```

CC TELEPHONE: (516) 742-4343
CC TELEFAX: (516) 742-4366
CC TELEX: 230 901 SANS UR
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 359 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 359 AA; 41425 MW; 698417 CN;

Query Match 4.4%; Score 155; DB 2; Length 359;
Best Local Similarity 32.8%; Pred. No. 1.67e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db 273 DEWAICQIRTTANFNPFV-NWFCGGLNHQVTHLFPNICHIIHYPOLENIKDVCOEF 331
QY 345 DNWEKQTKGTIDI-TCPP-WMDWFFGGLQFOLEHHLFRLPRGQLRKIAPLARDLCKKH 402
Db 332 GVEYKVY 338
QY 403 GMPYRSF 409

RESULT 8
ID US-08-307-382-2 STANDARD; PRT; 359 AA.

XX AC xxxxxx

Sequence 2, Application US/08307382

Sequence 2, Application US/08307382
Patent No. 5552306
GENERAL INFORMATION:
APPLICANT: Thomas, Terry L.
APPLICANT: Reddy, Avutu S.
APPLICANT: Nuccio, Michael
APPLICANT: Freyssinet, Georges L.
TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC
TITLE OF INVENTION: ACID BY A DELTA 6-DESATURASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/307,382
FILING DATE: 14-SEP-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/817,919
FILING DATE: 08-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: McNulty, William E.
REGISTRATION NUMBER: 22,606
REFERENCE/DOCKET NUMBER: 83832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids

CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 359 AA; 41425 MW; 698417 CN;

Query Match 4.4%; Score 155; DB 1; Length 359;
Best Local Similarity 32.8%; Pred. No. 1.67e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db 273 DEWAICQIRTTANFNPFV-NWFCGGLNHQVTHLFPNICHIIHYPOLENIKDVCOEF 331
QY 345 DNWEKQTKGTIDI-TCPP-WMDWFFGGLQFOLEHHLFRLPRGQLRKIAPLARDLCKKH 402
Db 332 GVEYKVY 338
QY 403 GMPYRSF 409

RESULT 9
ID US-08-478-727-2 STANDARD; PRT; 359 AA.

XX AC xxxxxx

Sequence 2, Application US/08478727

Sequence 2, Application US/08478727
Patent No. 5653068
GENERAL INFORMATION:
APPLICANT: Thomas, Terry L.
APPLICANT: Reddy, Avutu S.
APPLICANT: Nuccio, Michael
APPLICANT: Freyssinet, Georges L.
TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC
TITLE OF INVENTION: ACID BY A DELTA 6-DESATURASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,727
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/307,382
FILING DATE: 14-SEP-1994
APPLICATION NUMBER: US 07/817,919
FILING DATE: 08-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: McNulty, William E.
REGISTRATION NUMBER: 22,606
REFERENCE/DOCKET NUMBER: 83832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 359 AA; 41425 MW; 698417 CN;

CC SEQUENCE 359 AA; 41425 MW; 698417 CN;

Query Match 4.4%; Score 155; DB 1; Length 359;
Best Local Similarity 32.8%; Pred. No. 1.67e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db 273 DEWAICQIRTTANFNPFW-NWFCGGLNHQVTHLFPNICHHPQLENIKDYCQEF 331
QY 345 DNWFEKQTGTIDI-TCPP-WMDWFGGLQFQLEHLLFRLPRGQLRKTAPLARDLCKKH 402

Db 332 GVEYKVY 338
QY 403 GMPYRSF 409

RESULT 10
ID US-08-473-508-2 STANDARD; PRT: 359 AA.

XX
AC xxxxxx

Sequence 2, Application US/08473508

Sequence 2, Application US/08473508

Patent No. 5689050

GENERAL INFORMATION:

APPLICANT: Thomas, Terry L.

APPLICANT: Reddy, Avutu S.

APPLICANT: Nuccio, Michael

APPLICANT: Freyssinet, Georges L.

TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC

TITLE OF INVENTION: ACID BY A DELTA 6-DESATURASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: United States

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473,508

FILING DATE: 07-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/307,382

FILING DATE: 14-SEP-1994

APPLICATION NUMBER: US 07/817,919

FILING DATE: 08-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: McNulty, William E.

REGISTRATION NUMBER: 22,606

REFERENCE/DOCKET NUMBER: 8383Z

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 359 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 359 AA; 41425 MW; 698417 CN;

Query Match 4.4%; Score 155; DB 1; Length 359;
Best Local Similarity 32.8%; Pred. No. 1.67e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db 273 DEWAICQIRTTANFNPFW-NWFCGGLNHQVTHLFPNICHHPQLENIKDYCQEF 331
QY 345 DNWFEKQTGTIDI-TCPP-WMDWFGGLQFQLEHLLFRLPRGQLRKTAPLARDLCKKH 402

Db 332 GVEYKVY 338
QY 403 GMPYRSF 409

RESULT 11
ID US-08-801-972-5 STANDARD; PRT: 104 AA.

XX
AC xxxxxx

Sequence 5, Application US/08801972

Sequence 5, Application US/08801972

Patent No. 5831018

GENERAL INFORMATION:

APPLICANT: Goli, Surya K.

TITLE OF INVENTION: A NOVEL HUMAN CYTOCHROME

TITLE OF INVENTION: B5

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,972

FILING DATE: Filed Herewith

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0216 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 104 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1217655

SEQUENCE 104 AA; 11672 MW; 58506 CN;

Query Match 3.9%; Score 137; DB 2; Length 104;

Best Local Similarity 37.9%; Pred. No. 3.57e-02;

Matches 25; Conservative 11; Mismatches 29; Indels 1; Gaps 1;

Db 1 GRVYDITRFLSEHPGEEVLEQAGADATESFDYGHSPDAREMLKQYYIGDVHPNDLKP 60
QY 29 GRVYDCSRWAAEHGPGVEPLLSLAGQDVTDAFIAY-HPGTANRHLDPFTGYVYLKDFEV 87

Db 61 KGDGXD 66

QY 88 SEISKD 93

```

XX Sequence 2, Application US/08276099A
XX
XX Sequence 2, Application US/08276099A
XX Patent No. 5591825
XX GENERAL INFORMATION:
XX APPLICANT: Mcknight, Steven L
XX APPLICANT: Hou, Jinhao
XX TITLE OF INVENTION: INTERLEUKIN-4 SIGNAL TRANSDUCERS AND
XX TITLE OF INVENTION: BINDING ASSAYS
XX NUMBER OF SEQUENCES: 17
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
XX STREET: 4 Embarcadero Center, Suite 3400
XX CITY: San Francisco
XX STATE: California
XX COUNTRY: USA
XX ZIP: 94111-4187
XX COMPUTER READABLE FORM:
XX MEDIUM TYPE: Floppy disk
XX COMPUTER: IBM PC compatible
XX OPERATING SYSTEM: PC-DOS/MS-DOS
XX SOFTWARE: PatentIn Release #1.0, Version #1.25
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/08/276,099A
XX FILING DATE: 15-JUL-1994
XX CLASSIFICATION: 435
XX ATTORNEY/AGENT INFORMATION:
XX NAME: Osman, Richard Aron
XX REGISTRATION NUMBER: 36,627
XX REFERENCE/DOCKET NUMBER: A-59451-1/RAO
XX TELECOMMUNICATION INFORMATION:
XX TELEPHONE: (415) 781-1989
XX TELEFAX: (415) 398-3249
XX TELEX: 910 277299
XX INFORMATION FOR SEQ ID NO: 2:
XX SEQUENCE CHARACTERISTICS:
XX LENGTH: 847 amino acids
XX TYPE: amino acid
XX TOPOLOGY: linear
XX MOLECULE TYPE: protein
XX SEQUENCE 847 AA; 94134 MW; 3802900 CN;
XX
XX Query Match 2.9%; Score 102; DB 1; Length 847;
XX Best Local Similarity 21.18; Pred.No. 1.01e+01;
XX Matches 24; Conservative 36; Mismatches 44; Indels 10; Gaps
XX
Db 433 SEMDRVP-FYVAERVPWEKMCETLNLFKMAEVTGNRGGLP-EHFLFLAQKIFNDNLSLME 490
Qy 211 PDLQHLPVFAVSTRL-FNSITSVFYGRVLAFDEAVARLVSYQHWYTPYVMIFGRVNLFIQ 269
Db 491 AFQHSRVSWSQFNKEI-LLRGRTFFWQWFGVLDLTKRCLRSWSDRLIIGFIS 543
Qy 270 TFLDLLLRDVPDRALNMGIA-VFWTWFP--L-FVS-CLPN-WPERFGFVLIS 317
XX
RESULT 14
ID US-08-624-125-6 STANDARD; PRT; 175 AA.
XX
XX xxxxxx
XX
XX
XX
XX
XX Sequence 6, Application US/08624125
XX
XX Sequence 6, Application US/08624125
XX Patent No. 5744341
XX GENERAL INFORMATION:
XX APPLICANT: CUNNINGHAM JR., FRANCIS X.
XX APPLICANT: SUN, ZAIREN
XX TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
XX TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH G
XX NUMBER OF SEQUENCES: 21

```

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
CC ADDRESSEE: P.C.
CC STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CC CITY: ARLINGTON
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22202
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/624,125
CC FILING DATE: 29-MAR-1996
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: KELBER, STEVEN B.
CC REGISTRATION NUMBER: 30,073
CC REFERENCE/DOCKET NUMBER: 2747-063-27
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 703-413-3000
CC TELEFAX: 703-413-2220
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 175 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1/5 AA: 19867 MW: 149517 CN;

Query Match 2.8%; Score 97; DB 1; Length 175;
Best Local Similarity 28.8%; Pred. No. 2.17e+01;
Matches 23; Conservative 22; Mismatches 30; Indels 5; Gaps 5;
Db 37 HTPRKGVFELNDLFAV-VFAGVAITALIAGTAGVWPLQWICGTMVYGLLYFL-VHDGL 94
QY 99 NEMSRGIFKKGHHIMTFTVGAVNMAAI-VYGVLAESVGVHMLCGALLGLLTQAAAY 157
Db 95 V-HQRPWFHPIPRGYLKL 113
QY 158 VGHDSGHYQVMPTRGY-NRI 176

RESULT 15
US-08-096-623A-18 STANDARD; PRT; 176 AA.

AC xxxxxx

XX

DT

XX

DE

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/096,623A
CC FILING DATE: 22-JUL-1993
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/805,061
CC FILING DATE: 09-DEC-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/662,921
CC FILING DATE: 28-FEB-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/562,674
CC FILING DATE: 03-AUG-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/525,551
CC FILING DATE: 18-MAY-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/487,613
CC FILING DATE: 02-MAR-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Gamson, Edward P.
CC REGISTRATION NUMBER: 29,381
CC REFERENCE/DOCKET NUMBER: AMO-006.1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 655-1500
CC TELEFAX: (312) 655-1501
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 176 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 176 AA: 19990 MW: 151305 CN;

Query Match 2.8%; Score 97; DB 1; Length 176;
Best Local Similarity 28.8%; Pred. No. 2.17e+01;
Matches 23; Conservative 22; Mismatches 30; Indels 5; Gaps 5;

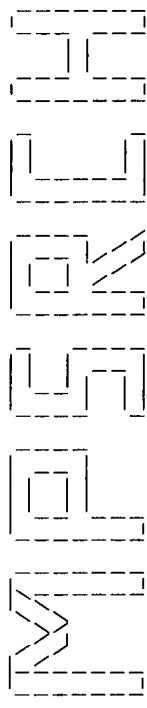
Db 38 HTPRKGVFELNDLFAV-VFAGVAITALIAGTAGVWPLQWICGTMVYGLLYFL-VHDGL 95
QY 99 NEMSRGIFKKGHHIMTFTVGAVNMAAI-VYGVLAESVGVHMLCGALLGLLTQAAAY 157

Db 96 V-HQRPWFHPIPRGYLKL 114

QY 158 VGHDSGHYQVMPTRGY-NRI 176

Search completed: Thu Feb 18 11:38:43 1999
Job time : 27 secs.

Sequence 18, Application US/08096623A
Sequence 18, Application US/08096623A
Patent No. 5684238
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L.
APPLICANT: Brinkhaus, Friedhelm L.
APPLICANT: Mukharji, Indrani
APPLICANT: Proffitt, John H.
APPLICANT: Yarger, James G.
APPLICANT: Yen, Hui-Che B.
TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESS: Welsh & Katz, Ltd.
STREET: 120 S. Riverside Plaza, 22nd Floor
CITY: Chicago
STATE: IL
COUNTRY: USA



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Feb 18 11:39:01 1999; MacPar time 23.20 Seconds
733.938 Million cell updates/sec
Linear output not generated.

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pap
Perfect Score: 3515
Sequence: 1 MEGSAKKYITAEEDLRHNSK.....LNSAPCPKRLGYGEYNTFHG 452

Scoring table: PAM 150
Gap 11

Searched: 321224 seqs, 37675139 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-pending
1:P9 2:U60 3:U7 4:U80 5:U81 6:U82 7:U83 8:U84A 9:U84B
10:U85 11:U86 12:U87 13:U88 14:U89 15:U90 16:U91 17:U92
18:NEWP 19:NEWU8 20:NEWU9

Statistics: Mean 36.598; Variance 169.128; scale 0.216

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Pred. No.	
t	Score	Query Match	Length	ID	Description	
1	3515	100.0	452	14	US-08-934- Sequence 27, Applicati	0.00e+00
2	2433	69.2	458	16	US-08-116- Sequence 5, Applicatio	1.16e-208
3	2133	60.7	448	2	US-60-110- Sequence 5, Applicatio	7.28e-181
4	2121	60.3	448	14	US-08-934- Sequence 5, Applicatio	9.39e-180
5	2108	60.0	446	13	US-08-834- Sequence 15, Applicati	1.50e-178
6	2108	60.0	445	14	US-08-956- Sequence 16, Applicati	1.50e-178
7	2108	60.0	446	13	US-08-833- Sequence 5, Applicatio	1.50e-178
8	2108	60.0	446	14	US-08-956- Sequence 16, Applicati	1.50e-178
9	1561	44.4	393	2	US-60-110- Sequence 4, Applicatio	4.80e-128
10	1557	44.3	252	14	US-08-956- Sequence 9, Applicatio	1.12e-127
11	1557	44.3	252	14	US-08-956- Sequence 9, Applicatio	1.12e-127
12	1557	44.3	252	13	US-08-834- Sequence 8, Applicatio	1.12e-127
13	1557	44.3	252	13	US-08-834- Sequence 7, Applicatio	1.12e-127
14	1457	41.5	286	2	US-60-110- Sequence 9, Applicatio	1.12e-127
15	1025	29.2	253	2	US-60-110- Sequence 11, Applicati	6.21e-79
16	897	25.5	162	2	US-60-110- Sequence 13, Applicati	2.65e-67
17	795	22.6	202	2	US-60-110- Sequence 2, Applicatio	4.45e-58
18	713	20.3	125	13	US-08-834- Sequence 8, Applicatio	1.07e-50
19	713	20.3	125	13	US-08-834- Sequence 9, Applicatio	1.07e-50
20	713	20.3	125	14	US-08-956- Sequence 10, Applicati	1.07e-50
21	713	20.3	125	14	US-08-956- Sequence 10, Applicati	1.07e-50

SQ SEQUENCE 452 AA; 51763 MW; 1074264 CN;

Query Match 100.0%; Score 3515; DB 14; Length 452;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MEGAKEYITAEEDLRHNSGDLWISIOGKYVDCSRWAAEHPPGGEVPLLSLAGQDVTDAF 60
 Qy 1 MEGAKEYITAEEDLRHNSGDLWISIOGKYVDCSRWAAEHPPGGEVPLLSLAGQDVTDAF 60
 Db 61 IAYHPGTAWRHLDPLFTCYGYYLKDFEVSISKDYRRLNEMSRSGIFFEKKGHHIMWTFVG 120
 Qy 61 IAYHPGTAWRHLDPLFTCYGYYLKDFEVSISKDYRRLNEMSRSGIFFEKKGHHIMWTFVG 120
 Db 121 VAVMAAIVYGVLSGVVHMLCGALLGLLWIAQAAVYGHDSGHYQVMPTRGYNRITOLI 180
 Qy 121 VAVMAAIVYGVLSGVVHMLCGALLGLLWIAQAAVYGHDSGHYQVMPTRGYNRITOLI 180
 Db 181 AGNLTGSIAMWKTNAHHLACNSLDYDPDLOHIPPFAVSTRFNSITSVFYGRVLKF 240
 Qy 181 AGNLTGSIAMWKTNAHHLACNSLDYDPDLOHIPPFAVSTRFNSITSVFYGRVLKF 240
 Db 241 DEVARFLVSYOHWTYYPMIFGRVNLFTQTFLLLLTRDVPDRALNMGIAVFWTFWPLF 300
 Qy 241 DEVARFLVSYOHWTYYPMIFGRVNLFTQTFLLLLTRDVPDRALNMGIAVFWTFWPLF 300
 Db 301 VSCLPNMPERFGLVLSFAVTAIOHVQFTLNHFSGDTYVGPCKGNWFEKOTKTIDITC 360
 Qy 301 VSCLPNMPERFGLVLSFAVTAIOHVQFTLNHFSGDTYVGPCKGNWFEKOTKTIDITC 360
 Db 361 PPMDWFFGGLOFQLEHLLFRLPRGQLRKIAPLARDLCKKHGMPYRSGFWDANVTI 420
 Qy 361 PPMDWFFGGLOFQLEHLLFRLPRGQLRKIAPLARDLCKKHGMPYRSGFWDANVTI 420
 Db 421 RTLDDAAVOARDLNSAPCKPKLGYGEAYNTHG 452
 Qy 421 RTLDDAAVOARDLNSAPCKPKLGYGEAYNTHG 452

RESULT 2
 ID US-09-116-639-5 STANDARD; PRT; 458 AA.
 XX
 AC xxxxxx
 XX
 DT
 XX
 DE
 XX
 Sequence 5, Application US/09116639
 Sequence 5, Application US/09116639
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Gorgone, Gina
 APPLICANT: Guebler, Karl J.
 APPLICANT: Corley, Neil C.
 APPLICANT: Baughn, Mariah R.
 TITLE OF INVENTION: HUMAN CYTOCHROMES
 CURRENT APPLICATION NUMBER: US/09/116.639
 CURRENT FILING DATE: 1998-07-16
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 5
 LENGTH: 458
 TYPE: PRT
 ORGANISM: HELIANTHUS ANNUUS
 FEATURE:
 OTHER INFORMATION: 1040729, GenBank
 SQ SEQUENCE 458 AA; 52231 MW; 1118207 CN;

Query Match 69.2%; Score 2433; DB 16; Length 458;
 Best Local Similarity 66.7%; Pred. No. 1.16e-208;
 Matches 300; Conservative 92; Mismatches 54; Indels 4; Gaps 4;

Db 13 ADGKXITSKELKNNPNDLWISILGKYVNTWAKHEHPGGDAPLINLAGQDVTDAFIA 72
 Qy 3 GEAKYITAEEDLRHNSGDLWISIOGKYVDCSRWAAEHPPGGEVPLLSLAGQDVTDAFIA 62
 Db 73 PHPGTAWRHLDPLFTCYGYYLKDFEVSISKDYRRLNEMSRSGIFFEKKGHHIMWTFVGVA 131
 Qy 63 YHPGTAWRHLDPLFTCYGYYLKDFEVSISKDYRRLNEMSRSGIFFEKKGHHIMWTFVGVA 122
 Db 132 LLLSACYGVLYSGFWIHLMSGAILGLAWMOIAYLGHDAHYQMMATRGWNKFAFIFIG 191
 Qy 123 VMMAAIVYGVLSGVVHMLCGALLGLLWIAQAAVYGHDSGHYQVMPTRGYNRITOLIAG 182
 Db 192 NCITGISTANWKTNAHHLACNSLDYDPDLOHLPMLAVSSKLFNSITSVFYGRQITFDP 251
 Qy 183 NILTGISTANWKTNAHHLACNSLDYDPDLOHIPPFAVSTRFNSITSVFYGRVLKFE 242
 Db 252 LARFVSVYOHWTYYPMIFGRVNLFTQTFLLLLTRDVPDRALNMGIAVFWTFWPLFVS 311
 Qy 243 VARFLVSYOHWTYYPMIFGRVNLFTQTFLLLLTRDVPDRALNMGIAVFWTFWPLFVS 302
 Db 312 RLPNMPERVAFLVSVFVTGQIQTNLNHFSGDVVYVGPCKGNWFEKOTKTIDILACS 371
 Qy 303 CLPNMPERFGLVLSFAVTAIOHVQFTLNHFSGDTYVGPCKGNWFEKOTKTIDITC 362
 Db 372 WMDWFFGGLOFQLEHLLFRLPRCHLRSIPICRELCKKYNLPYVSLSYD-ANVTTLKT 430
 Qy 363 WMDWFFGGLOFQLEHLLFRLPRGQLRKIAPLARDLCKKHGMPYRSGFWDANVTIRT 422
 Db 431 LRTAALQARDL-TNPAPONLAW-EAFNTHG 458
 Qy 423 LRDAAVQARDLNSAPCKPKLGYGEAYNTHG 452

RESULT 3
 ID US-60-110-784-5 STANDARD; PRT; 448 AA.
 XX
 AC xxxxxx
 XX
 DT
 XX
 DE
 XX
 Sequence 5, Application US/60110784
 Sequence 5, Application US/60110784
 GENERAL INFORMATION:
 APPLICANT: Kinney, Anthony J.
 APPLICANT: Cahoon, Edgar B.
 APPLICANT: Hitz, William D.
 APPLICANT: Cahoon, Rebecca E.
 TITLE OF INVENTION: MEMBRANE-BOUND DESATURASES
 FILE REFERENCE: BB-1264-P1
 CURRENT APPLICATION NUMBER: US/60/110.784
 CURRENT FILING DATE: 1998-12-03
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: Microsoft Word Version 7.0A
 SEQ ID NO 5
 LENGTH: 448
 TYPE: PRT
 ORGANISM: Borago officinalis
 SQ SEQUENCE 448 AA; 51634 MW; 1109964 CN;

Query Match 60.7%; Score 2133; DB 2; Length 448;
 Best Local Similarity 57.8%; Pred. No. 7.28e-181;
 Matches 262; Conservative 105; Mismatches 80; Indels 6; Gaps 5;

Db 1 MAAQIKYITSDKLNHDKPGDLWISIOGKYVDCSRWAAEHPPGGEVPLLSLAGQDVTDAF 60
 Qy 1 MEGAKEYITAEEDLRHNSGDLWISIOGKYVDCSRWAAEHPPGGEVPLLSLAGQDVTDAF 60
 Db 61 VAFHPASTWKNLDKFFTCYY-LKDYSEVSEVSKDYRRLNEMSRSGIFFEKKGHHIMWTFV 118
 Qy 61 IAYHPGTAWRHLDPLFTCYGYYLKDFEVSISKDYRRLNEMSRSGIFFEKKGHHIMWTFV 119

CC REGISTRATION NUMBER: 38,651
 CC REFERENCE/DOCKET NUMBER: CGAB-300.USA
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (415) 433-4150
 CC TELEFAX: (415) 433-8716
 CC TELEX: N/A
 CC INFORMATION FOR SEQ ID NO: 15:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 446 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: not relevant
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC SEQUENCE 446 AA; 51415 MW; 1102977 CN;
 SQ
 Query Match 60.0%; Score 2108; DB 13; Length 446;
 Best Local Similarity 57.8%; Pred. No. 1.50e-178;
 Matches 256; Conservative 104; Mismatches 78; Indels 5; Gaps 5;
 1 MAAQIKKIYTSDELKNDKPGDLWISIQGKAYDSDWKVDHPGSGFPLKSLAGQEVTDADF 60
 QY 1 MEGAARVITAEEDLRNRHSGDLWISIQGVYDCSRWAAEHGPGGEVPLLSLAGQDVTDAF 60
 Db 61 VAFHPASTWKNLDRFETGY-LKDYSEVSESVKVRKLVFEFSKMGLYDKKGH-IMEATLC 118
 QY 61 IAYHPTAWRHLDPLFTGYVYKDFEISEISKDYRLLNEMSRSGIFPKKGHHIMW-TFV 119
 Db 119 FIAMLFAMSVYGLVFCBGLVHLFSGCLMGFLWISIQGKAYDSDWKVDHPGSGFPLKSLAGQEVTDADF 60
 QY 120 GVAVMAAIVGVLAESVGVHMLCGALLGLLWIOAAVGHDSGHYQVMPTRGYNRIQTL 179
 Db 179 FAANCLSGISIGGWKWNHNAHIAACNSLEYDPLQYIPFLVYSSKFGSLTSHFYEKRLT 238
 QY 180 IAGNLTGISTAWKWTNNAHHLACNSLDYDPLQHPVFAVSTRLENSITSVFYGRVLK 239
 Db 239 FDSLRSFVSQHTWTFYPMCAARLNMVQSLIMLTAKRNVSYRAQELLCGLVFSIWPL 298
 QY 240 FDEVARFLVSQHTWTFYPMIFGRVNLFIQTFLLLTRDVPDRALNMGIAVFWTWFPL 299
 Db 299 LVSCLPNNGERIMFVIAISLVTMGOQVQFSLNHFSSSVYVGPKNWFEKOTDGLTDLIS 358
 QY 300 FVSCLPNNGERFGLVLSFAVTAIOHVQFTLNHFSGDITVGPKNWFEKOTKGTIDIT 359
 Db 359 CPPWMDWFGGLQFQIEHHLFPKMPRCNLRKISPVVIELCKKHNLPY-NYASFSEKANEMT 417
 QY 360 CPPWMDWFGGLQFQIEHHLFPKMPRCNLRKISPVVIELCKKHNLPY-NYASFSEKANEMT 417
 Db 418 LRTLRNTALQARDITK-PLPKNL 439
 QY 420 IRTLRAAQAQARDLNSAPCPKKL 442
 Db 442 IRTLRAAQAQARDLNSAPCPKKL 442
 RESULT 7
 ID US-08-956-985-16 STANDARD; PRT; 446 AA.
 XX
 AC
 DT
 TT
 XX
 DE
 XX
 CC Sequence 16, Application US/08956985
 CC GENERAL INFORMATION:
 CC APPLICANT: KNUZZON, DEBORAH
 CC TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
 CC OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
 CC NUMBER OF SEQUENCES: 28
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: RAE-VENTER LAW GROUP, P.C.
 CC STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039
 CC CITY: PALO ALTO
 CC STATE: CA
 CC COUNTRY: USA

CC ZIP: 94306
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/956,985
 CC FILING DATE: 24-OCT-1997
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/834,033
 CC FILING DATE: 11-APR-1997
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/833,610
 CC FILING DATE: 11-APR-1997
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: RAE-VENTER, BARBARA
 CC REGISTRATION NUMBER: 32,750
 CC REFERENCE/DOCKET NUMBER: CGNE.128.0105
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (650) 328-4400
 CC TELEFAX: (650) 328-4477
 CC TELEX: N/A
 CC INFORMATION FOR SEQ ID NO: 16:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 446 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: not relevant
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC SEQUENCE 446 AA; 51415 MW; 1102977 CN;
 SQ
 Query Match 60.0%; Score 2108; DB 14; Length 446;
 Best Local Similarity 57.8%; Pred. No. 1.50e-178;
 Matches 256; Conservative 104; Mismatches 78; Indels 5; Gaps 5;
 1 MAAQIKKIYTSDELKNDKPGDLWISIQGKAYDSDWKVDHPGSGFPLKSLAGQEVTDADF 60
 QY 1 MEGAARVITAEEDLRNRHSGDLWISIQGVYDCSRWAAEHGPGGEVPLLSLAGQDVTDAF 60
 Db 61 VAFHPASTWKNLDRFETGY-LKDYSEVSESVKVRKLVFEFSKMGLYDKKGH-IMEATLC 118
 QY 61 IAYHPTAWRHLDPLFTGYVYKDFEISEISKDYRLLNEMSRSGIFPKKGHHIMW-TFV 119
 Db 119 FIAMLFAMSVYGLVFCBGLVHLFSGCLMGFLWISIQGKAYDSDWKVDHPGSGFPLKSLAGQEVTDADF 60
 QY 120 GVAVMAAIVGVLAESVGVHMLCGALLGLLWIOAAVGHDSGHYQVMPTRGYNRIQTL 179
 Db 179 FAANCLSGISIGGWKWNHNAHIAACNSLEYDPLQYIPFLVYSSKFGSLTSHFYEKRLT 238
 QY 180 IAGNLTGISTAWKWTNNAHHLACNSLDYDPLQHPVFAVSTRLENSITSVFYGRVLK 239
 Db 239 FDSLRSFVSQHTWTFYPMCAARLNMVQSLIMLTAKRNVSYRAQELLCGLVFSIWPL 298
 QY 240 FDEVARFLVSQHTWTFYPMIFGRVNLFIQTFLLLTRDVPDRALNMGIAVFWTWFPL 299
 Db 299 LVSCLPNNGERIMFVIAISLVTMGOQVQFSLNHFSSSVYVGPKNWFEKOTDGLTDLIS 358
 QY 300 FVSCLPNNGERFGLVLSFAVTAIOHVQFTLNHFSGDITVGPKNWFEKOTKGTIDIT 359
 Db 359 CPPWMDWFGGLQFQIEHHLFPKMPRCNLRKISPVVIELCKKHNLPY-NYASFSEKANEMT 417
 QY 360 CPPWMDWFGGLQFQIEHHLFPKMPRCNLRKISPVVIELCKKHNLPY-NYASFSEKANEMT 417
 Db 418 LRTLRNTALQARDITK-PLPKNL 439
 QY 420 IRTLRAAQAQARDLNSAPCPKKL 442
 Db 442 IRTLRAAQAQARDLNSAPCPKKL 442
 RESULT 7
 ID US-08-833-610-5 STANDARD; PRT; 446 AA.
 XX

AC	XXXXXX	Sequence 5, Application US/08833610	Db	299	LVSCLPNWERIMFVIA	SVTGMQVQFSLNHFSSSVYVCKPKNWFEKQTDGTLDIS	355
XX		Sequence 5, Application US/08833610	XX				
DT		GENERAL INFORMATION:	DT				
XX		APPLICANT: KNUZON, DEBORAH	XX				
DE		APPLICANT: MURKJ, PRADIP	DE				
XX		APPLICANT: HUANG, YUNG-SHENG	XX				
CC		APPLICANT: THURMOND, JENNIFER	CC				
CC		APPLICANT: CHAUDHARY, SUNITA	CC				
CC		TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS	CC				
CC		TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS	CC				
CC		NUMBER OF SEQUENCES: 12	CC				
CC		CORRESPONDENCE ADDRESS:	CC				
CC		ADDRESSEE: RAE-VENTER LAW GROUP, P.C.	CC				
CC		STREET: 260 SHERIDAN AVE, P.O. BOX 60039	CC				
CC		CITY: PALO ALTO	CC				
CC		STATE: CALIFORNIA	CC				
CC		COUNTRY: USA	CC				
CC		ZIP: 94306	CC				
CC		COMPUTER READABLE FORM:	CC				
CC		MEDIUM TYPE: Floppy disk	CC				
CC		COMPUTER: IBM PC compatible	CC				
CC		OPERATING SYSTEM: PC-DOS/MS-DOS	CC				
CC		SOFTWARE: PatentIn Release #1.0, Version #1.30	CC				
CC		CURRENT APPLICATION DATA:	CC				
CC		APPLICATION NUMBER: US/08/833,610	CC				
CC		FILING DATE: 11-APR-1997	CC				
CC		CLASSIFICATION: 435	CC				
CC		ATTORNEY/AGENT INFORMATION:	CC				
CC		NAME: RAE-VENTER, BARBARA	CC				
CC		REGISTRATION NUMBER: 32,750	CC				
CC		REFERENCE/DOCKET NUMBER: CGNE.123.000S	CC				
CC		TELEPHONE: (650)328-4400	CC				
CC		TELEFAX: (650)328-4477	CC				
CC		TELEX: N/A	CC				
CC		INFORMATION FOR SEQ ID NO: 5:	CC				
CC		SEQUENCE CHARACTERISTICS:	CC				
CC		LENGTH: 446 amino acids	CC				
CC		TYPE: amino acid	CC				
CC		STRANDEDNESS: not relevant	CC				
CC		TOPOLOGY: linear	CC				
CC		MOLECULE TYPE: peptide	CC				
CC		SEQUENCE 446 AA; 51415 MW; 1102977 CN;	CC				
CC		Query Match 60.0%; Score 2108; DB 13; Length 446;	CC				
CC		Best Local Similarity 57.8%; Pred. No. 1.50e-178;	CC				
CC		Matches 256; Conservative 104; Mismatches 78; Indels 5; Gaps 5;	CC				
Db		1 MAAQIKYITSDKLNHDKPGDLWISIOGKAYDVSDWKDPGSGFPLKSLAGQEVTDFAF 60	Db				
Qy		1 MEGAKEYITAEIDRRHNKSGDLWISIOGKAYDVSDWKDPGSGFPLKSLAGQEVTDFAF 60	Qy				
Db		61 VAFHPASTWKNLDFCTGY-LKDYVSSEVSKYRKLVEFSKMGLYDKKGH-IMEATLC 118	Db				
Qy		61 IAYHPGTAWRLDPLFTGYIYKDFEVSISDYRLNEMSGRGEFEKKGHHIMW-TFV 119	Qy				
Db		119 FIAMLFAMSVYGVLFCEGVLVHLSGMLGFWLWISIOGKAYDVSDWKDPGSGFPLKSLAGQEVTDFAF 178	Db				
Qy		120 GVAVNMAAIVGVLAESVGVHMLGALLGLLWIAAYVGHDSGHYQVMPTRGYNRIQL 179	Qy				
Db		179 FAANCLSGISGWKWNNAHHIACNSLEYDPLDIYPLVYSSKFFGSLTSHFYEKRLT 238	Db				
Qy		180 IAGNLTGISAWKWTNAHHLACNSLDYDPLQHIPVFAVSTRLFNSTISVFGYGRVLK 239	Qy				
Db		239 FDSLRSREYVQHTFYPIMCAARLNVMVQSLIMLLTKRNVSYRAQELGCLVFSWYPL 298	Db				
Qy		240 FDEVARFLVSQHTWYTPVMIFGRVNLFIQTLELLLTTRDVPDRNLNMGIAVFWTFWFL 299	Qy				

Query Match

Best Local Similarity 74.6%; Pred. No. 1.12e-127;
Matches 188; Conservative 36; Mismatches 28; Indels 0; Gaps 0;

Db 1 GVLGYVLACTSVFAHQIAAALLGLLWISQAYIGHDSGHVYVMSKNSYNRFAQLLSGNCLT 60
QY 127 AIYGVGLASVGVHMLCGALLGLLWIAAYVGHDSGHYQVMPTRGYNRIQTLAGNILT 186
Db 61 GISIAWKKTHNAHHLACNSLDYDPDLQHIPVFAVSKFFSSLSRFDYDKLTFGPVARE 120
QY 187 GISIAWKKTHNAHHLACNSLDYDPDLQHIPVFAVSTRFNSITSFYGVLFKFEVARE 246
Db 121 LVSQHYFTYVNCVGRINLFQITFLLSKREVPPDRALNFAGILVFWTWPFLVSCLPN 180
QY 247 LVSQHYFTYVNCVGRINLFQITFLLSKREVPPDRALNFAGILVFWTWPFLVSCLPN 306
Db 181 WPERFFEVTSFTVTAHQIQTFLNHFADVYVGPPTGSDWFEKQAGTIDISCRSYMDW 240
QY 307 WPERFGEVLISFAVTAHQVQFTFLNHFSGDTYVGPFGKDNWFEKQKGTIDITCPPWMDW 366
Db 241 FFGGLQFQLEHH 252
367 FFGGLQFQLEHH 378

RESULT 11
ID US-08-956-985A-9 STANDARD; PRT: 252 AA.
XX
AC
XX
XX
DT
XX
DE Sequence 9, Application US/08956985A
XX
CC Sequence 9, Application US/08956985A
CC GENERAL INFORMATION:
CC APPLICANT: KNUTZON, DEBORAH
CC TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
CC TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS
CC NUMBER OF SEQUENCES: 28
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: RAE-VENTER LAW GROUP, P.C.
CC STREET: P.O. Box 60039
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94306-0039
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/956,985A
CC FILING DATE: 24-OCT-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/834,033
CC FILING DATE: 11-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/833,610
CC FILING DATE: 11-APR-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: RAE-VENTER, BARBARA
CC REGISTRATION NUMBER: 32,750
CC REFERENCE/DOCKET NUMBER: CGN128.01US
CC TELEPHONE: (650)328-4400
CC TELEFAX: (650)328-4477
CC TELEX: N/A
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 252 amino acids
CC TYPE: amino acid

CC STRANDEDNESS: not relevant
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
SQ SEQUENCE 252 AA; 28956 MW; 353818 CN;

Query Match 44.3%; Score 1557; DB 14; Length 252;
Best Local Similarity 74.6%; Pred. No. 1.12e-127;
Matches 188; Conservative 36; Mismatches 28; Indels 0; Gaps 0;

Db 1 GVLGYVLACTSVFAHQIAAALLGLLWISQAYIGHDSGHVYVMSKNSYNRFAQLLSGNCLT 60
QY 127 AIYGVGLASVGVHMLCGALLGLLWIAAYVGHDSGHYQVMPTRGYNRIQTLAGNILT 186
Db 61 GISIAWKKTHNAHHLACNSLDYDPDLQHIPVFAVSKFFSSLSRFDYDKLTFGPVARE 120
QY 187 GISIAWKKTHNAHHLACNSLDYDPDLQHIPVFAVSTRFNSITSFYGVLFKFEVARE 246
Db 121 LVSQHYFTYVNCVGRINLFQITFLLSKREVPPDRALNFAGILVFWTWPFLVSCLPN 180
QY 247 LVSQHYFTYVNCVGRINLFQITFLLSKREVPPDRALNFAGILVFWTWPFLVSCLPN 306
Db 181 WPERFFEVTSFTVTAHQIQTFLNHFADVYVGPPTGSDWFEKQAGTIDISCRSYMDW 240
QY 307 WPERFGEVLISFAVTAHQVQFTFLNHFSGDTYVGPFGKDNWFEKQKGTIDITCPPWMDW 366
Db 241 FFGGLQFQLEHH 252
367 FFGGLQFQLEHH 378

RESULT 12
ID US-08-834-033A-8 STANDARD; PRT: 252 AA.
XX
AC
XX
XX
DT
XX
DE Sequence 8, Application US/08834033A
XX
CC Sequence 8, Application US/08834033A
CC GENERAL INFORMATION:
CC APPLICANT: KNUTZON, DEBORAH
CC APPLICANT: MUKERJI, PRADIP
CC APPLICANT: HUANG, YUNG-SHENG
CC APPLICANT: THURMOND, JENNIFER
CC APPLICANT: CHAUDHARY, SUNITA
CC TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
CC TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
CC NUMBER OF SEQUENCES: 26
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.
CC STREET: 2001 FERRY BUILDING
CC CITY: SAN FRANCISCO
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: WORDPERFECT 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/834,033A
CC FILING DATE: 11-APR-1997
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: WARD, MICHAEL R.
CC REGISTRATION NUMBER: 38,651
CC REFERENCE/DOCKET NUMBER: CGAB-300.USA
CC TELEPHONE: (415) 433-4150
CC TELEFAX: (415) 433-8716
CC TELEX: N/A

CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 252 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: not relevant
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
SQ SEQUENCE 252 AA; 28956 MW; 353818 CN;

Query Match 44.3%; Score 1557; DB 13; Length 252;
Best Local Similarity 74.6%; Pred. No. 1.12e-127;
Matches 188; Conservative 36; Mismatches 28; Indels 0; Gaps 0;
Db 1 GVLGVLAQTSVFAHQIAAALLGLLWIOSAYIGHDSGHVIMSNKSYNRFQALLSGNCLT 60
Qy 127 AIYGVLAQTSVFAHQIAAALLGLLWIOSAYIGHDSGHVIMSNKSYNRFQALLSGNCLT 186
Db 61 GISIAWKTWTHNAHLACNSLDYDPDLQHIPVFAVSTKFFSLSLRSFYDRKLTGFPVARE 120
Qy 187 GISIAWKTWTHNAHLACNSLDYDPDLQHIPVFAVSTKFFSLSLRSFYDRKLTGFPVARE 246
Db 121 LVSQHTFYYPVNCGRINLFIQTLLFLSKREVDPDRALNFAVILVFTWTPLLVSCLPN 180
Qy 247 LVSQHTFYYPVNCGRINLFIQTLLFLSKREVDPDRALNFAVILVFTWTPLLVSCLPN 306
Db 181 WPERFFVFTFTALQHOFTLNHFAADVYVGPPTGSDWFEKQAACTIDISCRSYMDW 240
Qy 307 WPERFFVFTFTALQHOFTLNHFAADVYVGPPTGSDWFEKQAACTIDISCRSYMDW 366
Db 241 FFGGLQFQLEHH 252
Qy 367 FFGGLQFQLEHH 378

RESULT 13
ID US-08-834-655-7 STANDARD; PRT; 252 AA.

XX AC xxxxxx

Sequence 7, Application US/08834655

Sequence 7, Application US/08834655

GENERAL INFORMATION:

APPLICANT: KUTZON, DEBORAH

APPLICANT: MURKERT, PRADIP

APPLICANT: HUANG, YUNG-SHENG

APPLICANT: THURMOND, JENNIFER

APPLICANT: CHAUDHARY, SONITA

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS

TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: RAE-VENTER LAW GROUP, P.C.

STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/834,655

FILING DATE: 11-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: RAE-VENTER, BARBARA

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: CGNE.124.00US

CC.

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 328-4400
CC TELEFAX: (650) 328-4477
CC TELEX: N/A

CC INFORMATION FOR SEQ ID NO: 7:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 252 amino acids

CC TYPE: amino acid

CC STRANDEDNESS: not relevant

CC TOPOLOGY: linear

CC MOLECULE TYPE: peptide

SQ SEQUENCE 252 AA; 28956 MW; 353818 CN;

Query Match 44.3%; Score 1557; DB 13; Length 252;
Best Local Similarity 74.6%; Pred. No. 1.12e-127;
Matches 188; Conservative 36; Mismatches 28; Indels 0; Gaps 0;
Db 1 GVLGVLAQTSVFAHQIAAALLGLLWIOSAYIGHDSGHVIMSNKSYNRFQALLSGNCLT 60
Qy 127 AIYGVLAQTSVFAHQIAAALLGLLWIOSAYIGHDSGHVIMSNKSYNRFQALLSGNCLT 186
Db 61 GISIAWKTWTHNAHLACNSLDYDPDLQHIPVFAVSTKFFSLSLRSFYDRKLTGFPVARE 120
Qy 187 GISIAWKTWTHNAHLACNSLDYDPDLQHIPVFAVSTKFFSLSLRSFYDRKLTGFPVARE 246
Db 121 LVSQHTFYYPVNCGRINLFIQTLLFLSKREVDPDRALNFAVILVFTWTPLLVSCLPN 180
Qy 247 LVSQHTFYYPVNCGRINLFIQTLLFLSKREVDPDRALNFAVILVFTWTPLLVSCLPN 306
Db 181 WPERFFVFTFTALQHOFTLNHFAADVYVGPPTGSDWFEKQAACTIDISCRSYMDW 240
Qy 307 WPERFFVFTFTALQHOFTLNHFAADVYVGPPTGSDWFEKQAACTIDISCRSYMDW 366
Db 241 FFGGLQFQLEHH 252
Qy 367 FFGGLQFQLEHH 378

RESULT 14
ID US-60-110-784-9 STANDARD; PRT; 286 AA.

XX AC xxxxxx

Sequence 9, Application US/60110784

Sequence 9, Application US/60110784

GENERAL INFORMATION:

APPLICANT: Kinney, Anthony J.

APPLICANT: Cahoon, Edgar B.

APPLICANT: Hitz, William D.

APPLICANT: Cahoon, Rebecca E.

TITLE OF INVENTION: MEMBRANE-BOUND DESATURASES

FILE REFERENCE: BB-1264-P1

CURRENT APPLICATION NUMBER: US/60/110,784

CURRENT FILING DATE: 1998-12-03

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Microsoft Word Version 7.0A

SEQ ID NO 9

LENGTH: 286

TYPE: PRT

ORGANISM: Zea mays

SEQUENCE 286 AA; 32723 MW; 422187 CN;

Query Match 41.5%; Score 1457; DB 2; Length 286;
Best Local Similarity 62.0%; Pred. No. 1.77e-118;
Matches 176; Conservative 58; Mismatches 49; Indels 1; Gaps 1;

Db 1 IGFWIQSGWGHDSGHRRITGHPLLDVVQVLSNCLTGLSIWKNHNTHTHACNSL 60

Qy 148 LGLLWIOAAYVGHDSGHVQVMPTRGYNRITQLIAGNLTGISIAWKTWTHNAHLACNSL 207

[illegible]

WQSRFH
(TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 11:35:18 1999; MasPar time 18.91 Seconds
895.315 Million cell updates/sec
Similar output not generated.

Title: >US-08-934-254-27
Description: (1-452) from US08934254.ppe
Perfect Score: 3515
Sequence: 1 MEGEAKYITAEDLRHNKS.....LNSAPCPKILGYGEAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 116738 seqs, 37463448 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir58
1:pir1 2:pir2 3:pir3 4:pir4

Statistic: Mean 49.791; Variance 105.274; scale 0.473

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2433	69.2	458	2	S68358		hypothetical protein	0.00e+00
2	220	6.3	884	2	S66308		substrate-inducible n	4.72e-17
3	219	6.2	881	2	S25445		nitrate reductase (NA	6.84e-17
4	217	6.2	917	2	S35228		nitrate reductase (NA	1.44e-16
5	209	5.9	139	2	S4306		cytochrome b5 - commo	2.74e-15
6	203	5.8	120	2	S63052		cytochrome b5 - yeast	2.44e-14
7	203	5.8	137	2	S46307		cytochrome b5 - rice	2.44e-14
8	203	5.8	900	2	S47029		nitrate reductase (NA	2.44e-14
9	205	5.8	917	1	RDW0NH		nitrate reductase (NA	1.18e-14
10	200	5.7	911	1	RDTONH		nitrate reductase (NA	7.25e-14
11	202	5.7	918	2	A41667		nitrate reductase (NA	3.51e-14
12	199	5.7	926	1	RDSPNH		nitrate reductase (NA	1.04e-13
13	195	5.6	898	1	RDJNH		nitrate reductase (NA	3.07e-13
14	195	5.5	135	2	S49200		cytochrome b5 - commo	4.39e-13
15	193	5.5	135	1	CBRT5M		cytochrome b5, outer	9.00e-13
16	193	5.5	427	2	G70590		probable desA3, protei	9.00e-13
17	195	5.5	880	2	S72503		substrate-inducible n	4.39e-13
18	195	5.5	880	2	S66307		substrate-inducible n	4.39e-13
19	194	5.5	904	1	RDNTNT		nitrate reductase (NA	6.29e-13
20	193	5.5	904	1	RDNTNS		nitrate reductase (NA	9.00e-13
21	193	5.5	909	2	JN0665		nitrate reductase (NA	9.00e-13
22	192	5.5	982	2	S16292		nitrate reductase (NA	1.29e-12
23	189	5.4	368	2	S54809		linoleoyl-CoA desatur	3.75e-12

24	189	5.4	891	1	RDBHNP	nitrate reductase (NA	3.75e-12
25	188	5.3	74	2	A41085	nitrate reductase (NA	5.34e-12
26	187	5.3	501	2	S51160	nitrate reductase (NA	7.61e-12
27	186	5.3	912	1	RDBHNS	nitrate reductase (NA	1.08e-11
28	186	5.3	915	1	RDBHNS	nitrate reductase (NA	1.08e-11
29	183	5.2	617	2	S19254	nitrate reductase (NA	4.44e-11
30	183	5.2	916	2	S07554	nitrate reductase (NA	3.12e-11
31	182	5.2	920	2	S52301	nitrate reductase (NA	4.44e-11
32	174	5.0	87	2	S07959	cytochrome b5, hepati	7.20e-10
33	176	5.0	905	2	JN0803	nitrate reductase (NA	3.60e-10
34	171	4.9	98	2	JN0316	cytochrome b5, solubi	2.02e-09
35	172	4.9	133	2	S07964	cytochrome b5, hepati	1.43e-09
36	171	4.9	134	1	CBR85	cytochrome b5, hepati	2.02e-09
37	171	4.9	134	1	CBR05	cytochrome b5 - orise	2.02e-09
38	170	4.8	97	2	D24211	cytochrome b5, erythr	2.85e-09
39	170	4.8	98	1	CBH05E	cytochrome b5, erythr	2.85e-09
40	170	4.8	98	2	JC5783	soluble cytochrome b5	2.85e-09
41	170	4.8	133	2	C24211	cytochrome b5, hepati	2.85e-09
42	170	4.8	134	1	CBH05	cytochrome b5, micros	2.85e-09
43	170	4.8	134	2	JC5782	membrane-bound cytoch	2.85e-09
44	167	4.8	134	1	CBR05	cytochrome b5, micros	7.93e-09
45	168	4.8	147	2	S67453	hypothetical protein	5.64e-09

ALIGNMENTS

RESULT 1
ENTRY #type complete
TITLE hypothetical protein - common sunflower
ORGANISM #formal_name Helianthus annuus #common_name common sunflower
DATE 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 10-Sep-1997
ACCESSION S68358
REFERENCE S68358
#authors Sperling, P.; Schmidt, H.; Heinz, E.
#journal Eur. J. Biochem. (1995) 232:798-805
#title A cytochrome-b(5)-containing fusion protein similar to plant acyl lipid desaturases.
#accession S68358
#status preliminary
#molecule_type mRNA
#residues 1-458 #label SPE
#cross-references EMBL:X87143; NID:g1040738; PID:g1040729
CLASSIFICATION #superfamily cytochrome b5 core homology
FEATURE
16-90 #domain cytochrome b5 core homology #label CB5
SUMMARY #length 458 #molecular-weight 52231 #checksum 8719

Query Match	69.2%	Score	2433	DB 2	Length	458	
Best Local Similarity	66.7%	Pred. No.	0.00e+00				
Matches	300	Conservative	92	Mismatches	4	Gaps	4
Db	13	ADGKKYITSEKELKNNPNDLWISILGKYNYVTEWAKEHPGGADPLINLAGQDVTDAFIA 72					
QY	3	GKAAKYITAEEDLRHNKSGDLWISIQGYDCSRWAAEHPGGEVPLSLAGQDVTDAFIA 62					
Db	73	FHPGTAWKHLDTGYH-LKDYQVSDISRDYRKLAFAKAGMEPKKGHGYISLCFVS 131					
QY	63	YHPGTAWRHLDPLFTGYTYKDFEVSISKDYRRLNEMSRSGIFKEKGGHHNIWTFVGVA 122					
Db	132	LLISACVGYVSGSFWIHMGLGAILGLAWMQTAYLGHDAGHYQMATTGWNKFAFIFIG 191					
QY	123	VMAAIVGYVLASESVGVHMLCGALLGLLWIAQAAVYGHDSGHVQVMPTRGYNRIITQIAG 182					
Db	192	NCITGISIAWKWTHNAHIAACNSLDYDPDLQHLPLMVAVSSKLFNSITSVFYGRQRTFDP 251					
QY	183	NILTGISIAWKWTHNAHIAACNSLDYDPDLQHIPVFAVSTRLFNSTSVFYGRVCLKFDE 242					
Db	252	LARFFVSYQHYLYPYPCMVARNVLYIQTILLISKRKIPDRGLNLTGLFTWTFPLVLS 311					
QY	243	VARFLVSYQHYTYPYPIFYGRVNLFTQTILLTRDVPDRALNMGIAVFWTFPLVLS 302					
Db	312	RLPNWPERVAFLVSCVTGIQIOTFLNHFSGDVYVGGPPKGNWFEKQTRGTIDTACSS 371					

```

QY 303 CLPNWPERFGLVLSFAVTAIQHVQFTLNHFSGDITYVGPCKGNWFEKOTKGTIDITCPP 362
Db 372 WMDWFFGGLQFQLEHLHFLPRLPCHRSLSPICRELCKXKYNLPYVLSFYD-ANVTTLKF 430
QY 363 WMDWFFGGLQFQLEHLHFLPRLPGQLRKIAPLARDLCKKHGMPYRSGFGWDDANVTIRT 422
Db 431 LRTAAIQARDL-TNPAPONLAW-EARNTHG 458
QY 423 LRDAAVQARDLNSAPCPKILGYGEAYNTHG 452

RESULT 2
ENTRY substrate-inducible nitrate reductase 2 - soybean (fragment)
TITLE #formal_name Glycine max #common_name soybean
ORGANISM 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change
DATE 20-Mar-1998
ACCESSIONS S66308
REFERENCE S66307
AUTHORS Wu, S.; Lu, Q.; Kriz, A.L.; Harper, J.E.
JOURNAL Plant Mol. Biol. (1995) 29:491-506
TITLE Identification of cDNA clones corresponding to two inducible
#title nitrate reductase genes in soybean: analysis in wild-type
#accession S66308
#status preliminary
#molecule_type mRNA
#residues 1-884 #label WUS
#cross-references EMBL:U13987
#note the authors translated the codon CTC for residue 7 as
#note ile
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core
#domain cytochrome b5 core homology #label CB5\
#domain cytochrome-b5 reductase homology #label CBR
#length 884 #checksum 3299
FEATURE 59-451
511-585 #domain molybdopterin-binding domain homology #label
635-884 PCO\
SUMMARY #domain cytochrome b5 core homology #label CB5\
#domain cytochrome-b5 reductase homology #label CBR
#length 884 #checksum 3299
Query Match 6.3%; Score 220; DB 2; Length 884;
Best Local Similarity 41.1%; Pred. No. 4.72e-17;
Matches 30; Conservative 16; Mismatches 26; Indels 1; Gaps 1;
Db 507 MNTTSKTYTMSV-VRRHNADSAWIIIVGHVYDCTFLKDHGGTDSILINAGTCTEEF 565
1 MEGEAKYITAEPLDRHNKSGDLWISIQGVKVDSCRWAAEHFPGVEYPLLSLAGQVDYDAF 60
Db 566 EATHSOKAKOMLE 578
QY 61 IAYHPTAWRHL 73

RESULT 3
ENTRY S25445 #type complete
TITLE nitrate reductase (NADH) (EC 1.6.6.1) 1 - kidney bean
ORGANISM #formal_name Phaseolus vulgaris #common_name kidney bean
DATE 22-Nov-1993 #sequence_revision 08-Sep-1995 #text_change
ACCESSIONS S25445
REFERENCE S25445
AUTHORS Hoff, T.; Stummann, B.M.; Henningsen, K.W.
JOURNAL Physiol. Plantarum (1991) 82:197-204
TITLE Cloning and expression of a gene encoding a root specific
#title nitrate reductase in bean (Phaseolus vulgaris).
#accession S25445
#molecule_type DNA
#residues 1-881 #label HOF
#cross-references EMBL:X53603; NID:g21018; PID:g21019
#experimental_source cv. Saxonia
GENETICS

```



```

ACCESSIONS S63052; S47919; S67347
REFERENCE S63047
#authors De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.; Valle, G.
#submission submitted to the Protein Sequence Database, April 1996
#accession S63052
##molecule_type DNA
##residues 1-120 #label DEA
##cross-references EMBL:D71387; NID:gi302031; PID:g239931; PID:gi302032; MIPS:YNLI11C
##experimental_source strain S288C
REFERENCE S47919
#authors Truan, G.; Epinat, J.C.; Rougeulle, C.; Cullin, C.; Pompon, D.
#journal Gene (1994) 149:123-127
#title Cloning and characterization of a yeast cytochrome b(5)-encoding gene which suppresses ketoconazole hypersensitivity in a NADPH-P-450 reductase-deficient strain.
#accession S47919
##molecule_type DNA
##residues 1-16,'O',18-120 #label TRU
##cross-references EMBL:I22494; NID:g431761; PID:g431762
REFERENCE S67327
#authors d'Angelo, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.; Valle, G.
#submission submitted to the EMBL Data Library, February 1996
#description The DNA sequence of cosmid 14-13b from chromosome XIV of Saccharomyces cerevisiae reveals an unusually high number of overlapping ORFs.
#accession S67347
##molecule_type DNA
##residues 1-120 #label DAN
##cross-references EMBL:Z69382; NID:gil183941; PID:e221832; PID:gil183962
GENETICS
#gene SGD:CYB5
#map_position 14L
FUNCTION electron transfer
CLASSIFICATION superfamily cytochrome b5; cytochrome b5 core homology
KEYWORDS electron transfer; heme; transmembrane protein
FEATURE
2-77 #domain cytochrome b5 core homology #label CB5\
102-118 #domain transmembrane #status predicted #label TMM
SUMMARY #length 120 #molecular-weight 13297 #checksum 2403
Query Match 5.8%; Score 203; DB 2; Length 120;
Best Local Similarity 38.7%; Pred. No. 2.44e-14;
Matches 29; Conservative 19; Mismatches 25; Indels 2; Gaps 2;
Db 2 PKVY-SYQVEAENGPENFWIIDDKYVDYSQFDEHPGGEIMDLGGQDTSFVDIG 60
:| : :: | : :: | : || | : || : | : || : | : || : | : || : | :
Qy 5 AKKVITAEELRRHKSGDLWISIQGVKDYDCSRWAAEHPPGGEVTLSSLAGQDVDTAFAY- 63
Db 61 HSDEALRLKLGLYIG 75
:| : || | : || : | : || : | : || : | : || : | : || : | :
Qy 64 HPGTAWRHLDPLFTG 78
RESULT 7
ENTRY cytochrome b5 - rice
TITLE #type complete
ORGANISM #formal_name Oryza sativa #common_name rice
DATE 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 08-Sep-1997
ACCESSIONS S46307; S38582
REFERENCE S46306
#authors Smith, M.A.; Stobart, A.K.; Shewry, P.R.; Napier, J.A.
#journal Plant Mol. Biol. (1994) 25:527-537
#title Tobacco cytochrome b(5): cDNA isolation, expression analysis and in vitro protein targeting.
#accession S46307

```


J. Biol. Chem. (1991) 266:23542-23547
The sequence of squash NADH:nitrate reductase and its relationship to the sequences of other flavoprotein oxidoreductases. A family of flavoprotein pyridine nucleotide cytochrome reductases.

#cross-references MUID:92084635
#accession A41667
#molecule_type mRNA
#residues 1-918 #label HYD
#cross-references GB:M3154; NID:G167498; PID:G167499
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reductase homology; molybdopterin-binding domain homology; chromoprotein; electron transfer; FAD; flavoprotein; heme; homodimer; iron; metalloprotein; molybdenum; molybdopterin; NAD; nitrate assimilation; oxidoreductase

KEYWORDS
FEATURE 91-483
543-617 #domain cytochrome b5 core homology #label CB5\
668-918 #domain cytochrome-b5 reductase homology #label CBR\
195 #binding_site molybdopterin (Cys) (covalent) #status predicted\
434 #disulfide_bonds interchain #status predicted\
578,601 #binding_site heme iron (His) (axial ligands) #status predicted\
733,890 #binding_site NAD(P) (Lys, Cys) #status predicted\
773 #binding_site FAD (Tyr) #status predicted
SUMMARY #length 918 #molecular_weight 103363 #checksum 8142

Query Match 5.7%; Score 202; DB 2; Length 918;
Best Local Similarity 37.0%; Pred. No. 3.51e-14;
Matches 27; Conservative 18; Mismatches 27; Indels 1; Gaps 1;

Db 539 MNTASNTY-THSEVKKNSQSAWIIVGHVYDCTRELKDPGGSDILNAGTDCTEEF 597
QY 1 MEGEAKYITAEPLRRHNSGDLWISIQGVYDCSRWAEHPGGEVPLSLAGQDVTDFA 60

Db 598 DAHSKAKKLE 610
QY 61 IAYHPTAWRHL 73

RESULT 12
ENTRY RDSPNH #type complete
TITLE nitrate reductase (NADH) (EC 1.6.6.1) - spinach
ORGANISM #formal_name Spinacia oleracea #common_name spinach
DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Mar-1998
ACCESSIONS S11868; PQ0694; PQ0695; S62777; A45589; B45589; S68383
REFERENCE S11868
#authors Prosser, I.M.; Lazarus, C.M.
#journal Plant Mol. Biol. (1990) 15:187-190
#title Nucleotide sequence of a spinach nitrate reductase cDNA.
#cross-references MUID:91355858
#accession S11868
#molecule_type mRNA
#residues 1-926 #label PRO
#cross-references GB:M32600; NID:G170118; PID:G170119
REFERENCE PQ0694
#authors Shiraiishi, N.; Kubo, Y.; Takeba, G.; Kiyota, S.; Sakano, K.; Nakagawa, H.
#journal Plant Cell Physiol. (1991) 32:1031-1038
#title Sequence analysis of cloned cDNA and proteolytic fragments for nitrate reductase from Spinacia oleracea L.
#accession PQ0694
#molecule_type mRNA
#residues 287-926 #label SHI
#experimental_source cv. Hoyo
#accession PQ0695
#molecule_type protein
#residues 541-550; 659-667 #label SH2
REFERENCE S62777

J. Biol. Chem. (1991) 266:23542-23547
The sequence of squash NADH:nitrate reductase and its relationship to the sequences of other flavoprotein oxidoreductases. A family of flavoprotein pyridine nucleotide cytochrome reductases.

#cross-references MUID:92084635
#accession A41667
#molecule_type mRNA
#residues 1-918 #label HYD
#cross-references GB:M3154; NID:G167498; PID:G167499
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reductase homology; molybdopterin-binding domain homology; chromoprotein; electron transfer; FAD; flavoprotein; heme; homodimer; iron; metalloprotein; molybdenum; molybdopterin; NAD; nitrate assimilation; oxidoreductase

KEYWORDS
FEATURE 91-483
543-617 #domain cytochrome b5 core homology #label CB5\
668-918 #domain cytochrome-b5 reductase homology #label CBR\
195 #binding_site molybdopterin (Cys) (covalent) #status predicted\
434 #disulfide_bonds interchain #status predicted\
578,601 #binding_site heme iron (His) (axial ligands) #status predicted\
733,890 #binding_site NAD(P) (Lys, Cys) #status predicted\
773 #binding_site FAD (Tyr) #status predicted
SUMMARY #length 918 #molecular_weight 103363 #checksum 8142

Query Match 5.7%; Score 202; DB 2; Length 918;
Best Local Similarity 37.0%; Pred. No. 3.51e-14;
Matches 27; Conservative 18; Mismatches 27; Indels 1; Gaps 1;

Db 539 MNTASNTY-THSEVKKNSQSAWIIVGHVYDCTRELKDPGGSDILNAGTDCTEEF 597
QY 1 MEGEAKYITAEPLRRHNSGDLWISIQGVYDCSRWAEHPGGEVPLSLAGQDVTDFA 60

Db 598 DAHSKAKKLE 610
QY 61 IAYHPTAWRHL 73

RESULT 12
ENTRY RDSPNH #type complete
TITLE nitrate reductase (NADH) (EC 1.6.6.1) - spinach
ORGANISM #formal_name Spinacia oleracea #common_name spinach
DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Mar-1998
ACCESSIONS S11868; PQ0694; PQ0695; S62777; A45589; B45589; S68383
REFERENCE S11868
#authors Prosser, I.M.; Lazarus, C.M.
#journal Plant Mol. Biol. (1990) 15:187-190
#title Nucleotide sequence of a spinach nitrate reductase cDNA.
#cross-references MUID:91355858
#accession S11868
#molecule_type mRNA
#residues 1-926 #label PRO
#cross-references GB:M32600; NID:G170118; PID:G170119
REFERENCE PQ0694
#authors Shiraiishi, N.; Kubo, Y.; Takeba, G.; Kiyota, S.; Sakano, K.; Nakagawa, H.
#journal Plant Cell Physiol. (1991) 32:1031-1038
#title Sequence analysis of cloned cDNA and proteolytic fragments for nitrate reductase from Spinacia oleracea L.
#accession PQ0694
#molecule_type mRNA
#residues 287-926 #label SHI
#experimental_source cv. Hoyo
#accession PQ0695
#molecule_type protein
#residues 541-550; 659-667 #label SH2
REFERENCE S62777

J. Biol. Chem. (1991) 266:23542-23547
The sequence of squash NADH:nitrate reductase and its relationship to the sequences of other flavoprotein oxidoreductases. A family of flavoprotein pyridine nucleotide cytochrome reductases.

#cross-references MUID:92084635
#accession A41667
#molecule_type mRNA
#residues 1-918 #label HYD
#cross-references GB:M3154; NID:G167498; PID:G167499
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reductase homology; molybdopterin-binding domain homology; chromoprotein; electron transfer; FAD; flavoprotein; heme; homodimer; iron; metalloprotein; molybdenum; molybdopterin; NAD; nitrate assimilation; oxidoreductase

KEYWORDS
FEATURE 91-483
543-617 #domain cytochrome b5 core homology #label CB5\
668-918 #domain cytochrome-b5 reductase homology #label CBR\
195 #binding_site molybdopterin (Cys) (covalent) #status predicted\
434 #disulfide_bonds interchain #status predicted\
578,601 #binding_site heme iron (His) (axial ligands) #status predicted\
733,890 #binding_site NAD(P) (Lys, Cys) #status predicted\
773 #binding_site FAD (Tyr) #status predicted
SUMMARY #length 918 #molecular_weight 103363 #checksum 8142

Query Match 5.7%; Score 202; DB 2; Length 918;
Best Local Similarity 37.0%; Pred. No. 3.51e-14;
Matches 27; Conservative 18; Mismatches 27; Indels 1; Gaps 1;

Db 539 MNTASNTY-THSEVKKNSQSAWIIVGHVYDCTRELKDPGGSDILNAGTDCTEEF 597
QY 1 MEGEAKYITAEPLRRHNSGDLWISIQGVYDCSRWAEHPGGEVPLSLAGQDVTDFA 60

Db 598 DAHSKAKKLE 610
QY 61 IAYHPTAWRHL 73

RESULT 12
ENTRY RDSPNH #type complete
TITLE nitrate reductase (NADH) (EC 1.6.6.1) - spinach
ORGANISM #formal_name Spinacia oleracea #common_name spinach
DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Mar-1998
ACCESSIONS S11868; PQ0694; PQ0695; S62777; A45589; B45589; S68383
REFERENCE S11868
#authors Prosser, I.M.; Lazarus, C.M.
#journal Plant Mol. Biol. (1990) 15:187-190
#title Nucleotide sequence of a spinach nitrate reductase cDNA.
#cross-references MUID:91355858
#accession S11868
#molecule_type mRNA
#residues 1-926 #label PRO
#cross-references GB:M32600; NID:G170118; PID:G170119
REFERENCE PQ0694
#authors Shiraiishi, N.; Kubo, Y.; Takeba, G.; Kiyota, S.; Sakano, K.; Nakagawa, H.
#journal Plant Cell Physiol. (1991) 32:1031-1038
#title Sequence analysis of cloned cDNA and proteolytic fragments for nitrate reductase from Spinacia oleracea L.
#accession PQ0694
#molecule_type mRNA
#residues 287-926 #label SHI
#experimental_source cv. Hoyo
#accession PQ0695
#molecule_type protein
#residues 541-550; 659-667 #label SH2
REFERENCE S62777

J. Biol. Chem. (1991) 266:23542-23547
The sequence of squash NADH:nitrate reductase and its relationship to the sequences of other flavoprotein oxidoreductases. A family of flavoprotein pyridine nucleotide cytochrome reductases.

#cross-references MUID:92084635
#accession A41667
#molecule_type mRNA
#residues 1-918 #label HYD
#cross-references GB:M3154; NID:G167498; PID:G167499
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reductase homology; molybdopterin-binding domain homology; chromoprotein; electron transfer; FAD; flavoprotein; heme; homodimer; iron; metalloprotein; molybdenum; molybdopterin; NAD; nitrate assimilation; oxidoreductase

KEYWORDS
FEATURE 91-483
543-617 #domain cytochrome b5 core homology #label CB5\
668-918 #domain cytochrome-b5 reductase homology #label CBR\
195 #binding_site molybdopterin (Cys) (covalent) #status predicted\
434 #disulfide_bonds interchain #status predicted\
578,601 #binding_site heme iron (His) (axial ligands) #status predicted\
733,890 #binding_site NAD(P) (Lys, Cys) #status predicted\
773 #binding_site FAD (Tyr) #status predicted
SUMMARY #length 918 #molecular_weight 103363 #checksum 8142

Query Match

```

Quinn, G.B.; Trimball, A.J.; Prosser, I.M.; Barber, M.J.  
#authors  
#journal Arch. Biochem. Biophys. (1996) 327:151-160  
#title Spectroscopic and kinetic properties of a recombinant form of  
the flavin domain of spinach NADH:nitrate reductase.  
#accession S62777  
##molecule_type protein  
##residues 659-673;722-731;758-766;789-798;802-814;829-833;846-852;  
853-858;859-864;874-879;904-914;916-926 ##label QUI  
REFERENCE  
A45589  
#authors Fido, R.J.  
#journal Phytochemistry (1991) 30:3519-3523  
#title Isolation and partial amino acid sequence of domains of  
nitrate reductase from spinach.  
#accession A45589  
##molecule_type protein  
##residues 659-701 ##label FID  
##note sequence extracted from NCBI backbone (NCBIP:79483)  
#accession B45589  
##molecule_type protein  
##residues 'A',533-564 ##label FI2  
##note sequence extracted from NCBI backbone (NCBIP:79485)  
REFERENCE  
S68382  
#authors Douglas, P.; Morrice, N.; Mackintosh, C.  
#journal FEBS Lett. (1995) 377:113-117  
#title Identification of a regulatory phosphorylation site in the  
hinge I region of nitrate reductase from spinach (Spinacea  
oleracea) leaves.  
#accession S68383  
##molecule_type protein  
##residues 25-33;46-60;541-552 ##label DOU  
COMPLEX homodimer  
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core  
homology; cytochrome-b5 reductase homology;  
molybdopterin-binding domain homology  
KEYWORDS chromoprotein; electron transfer; FAD; flavoprotein; heme;  
homodimer; iron; metalloprotein; molybdenum; molybdopterin;  
NAD; nitrate assimilation; oxidoreductase; phosphoprotein  
FEATURE  
100-492 #domain molybdopterin-binding domain homology #label  
PCO\N  
551-625 #domain cytochrome b5 core homology #label CB5\  
677-926 #domain cytochrome-b5 reductase homology #label CBR\  
204 #binding_site molybdopterin (Cys) (covalent) #status  
predicted\  
443 #disulfide_bonds interchain #status predicted\  
543 #binding_site phosphate (Ser) (covalent) (by nitrate  
reductase-inactivating kinase) #status experimental\  
586,609 #binding_site heme iron (His) (axial ligands) #status  
predicted\  
741,898 #binding_site NAD (lys, Cys) #status predicted\  
781 #binding_site FAD (Tyr) #status predicted  
SUMMARY #length 926 #molecular_weight 103970 #checksum 1366  
  
Query Match 5.7%; Score 199; DB 1; Length 926;  
Best Local Similarity 37.0%; Pred. No.1.04e-13;  
Matches 27; Conservative 19; Mismatches 26; Indels 1; Gaps 1;  
  
Db 547 MNITSKWSMSF-VKKHNTADSAWITVHGNYNATRFLLKHDPGSDSLINAGDCTEEF 605  
QY ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||  
1 MEGEARKYTIEDLRHRNKSGDLWISIOGVDCSRNAEHGPGEVPLLGLAGDVTDAF 60  
  
Db 606 DAIHSDKAKRLLE 618  
QY |::|::|:  
61 IAYHPCTAWRHLD 73  
  
RESULT 13  
ENTRY RDBJNH #type complete  
TITLE nitrate reductase (NAD(P>H) (EC 1.6.6.2) - European white  
birch  
#formal_name Betula pendula #common_name European white birch  
DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change  
20-Mar-1998

```

Search completed: Thu Feb 18 11:35:47 1999
Job time : 29 secs.

W P S R E H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Thu Feb 18 17:54:20 1999; MasPar time 1707.01 Seconds
1334.208 Million cell updates/sec
Near output not generated.

Title: >US-08-934-254-26
Description: (1-1702) from US08934254.seq
Perfect Score: 1702 1 CCCCCAAATTTTCATTGTT.....TTTTTGTAATAAAAAAAA 1702
N.A. Sequence: Comp: GGGGTTTTTAAAGTACAA.....AAAAACCATTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2077918 seqs, 66906966 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-pending
1:P9 2:U6000 3:U6001 4:U6002A 5:U6002B 6:U6003A 7:U6003B
8:U6004A 9:U6004B 10:U6005 11:U6006 12:U6007 13:U6008A
14:U6008B 15:U6008C 16:U6009A 17:U6009B 18:U6010A
19:U6010B 20:U7 21:U80 22:U81A 23:U81B 24:U81C 25:U82A
26:U82B 27:U82C 28:U83A 29:U83B 30:U84A 31:U84B 32:U84C
33:U85 34:U86 35:U87A 36:U87B 37:U87C 38:U88A 39:U88B
40:U88C 41:U88D 42:U89A 43:U89B 44:U89C 45:U89D 46:U89E
47:U90A 48:U90B 49:U90C 50:U91A 51:U91B 52:U92 53:NEWP
54:NEW08 55:NEW09

Statistics: Mean 10.769; Variance 3.383; scale 3.183

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1702	100.0	1702	43	US-08-934- Sequence 26, Applicati	0.00e+00
2	172	10.1	1684	39	US-08-831- Sequence 1, Applicatio	4.24e-178
3	172	10.1	1684	39	US-08-831- Sequence 1, Applicatio	4.24e-178
4	172	10.1	1685	43	US-08-934- Sequence 4, Applicatio	4.24e-178
5	116	6.8	291	14	US-60-086- Sequence 1966, Applic	2.32e-106
6	110	6.5	276	14	US-60-086- Sequence 2042, Applic	7.87e-99
7	103	6.1	267	15	US-60-089- Sequence 1341, Applic	4.23e-90
8	98	5.8	298	13	US-60-080- Sequence 1738, Applic	6.63e-84
9	98	5.8	304	15	US-60-080- Sequence 1177, Applic	6.63e-84
10	97	5.7	578	34	US-08-624- Sequence 14, Applicati	1.14e-82
11	97	5.7	578	33	US-08-569- Sequence 14, Applicati	1.14e-82
12	90	5.3	266	14	US-60-086- Sequence 3256, Applic	4.52e-74

ID	US-08-934-254-26	STANDARD; DNA; UNC; 1702 BP.	ALIGNMENTS
13	91	5.3	7218 31 US-08-466- Sequence 14, Applicati
14	89	5.2	578 33 US-08-569- Sequence 14, Applicati
15	89	5.2	578 34 US-08-624- Sequence 14, Applicati
16	86	5.1	526 15 US-60-089- Sequence 1148, Applic
17	85	5.0	250 14 US-60-085- Sequence 3638, Applic
18	85	5.0	250 15 US-60-089- Sequence 569, Applicat
19	84	4.9	284 15 US-60-089- Sequence 503, Applicat
20	84	4.9	284 13 US-60-080- Sequence 2485, Applic
21	75	4.4	383 31 US-08-446- Sequence 8, Applicatio
22	75	4.4	383 28 US-08-311- Sequence 8, Applicatio
23	75	4.4	383 31 US-08-446- Sequence 8, Applicatio
24	75	4.4	383 31 US-08-451- Sequence 8, Applicatio
25	75	4.4	383 31 US-08-446- Sequence 8, Applicatio
26	75	4.4	383 31 US-08-446- Sequence 8, Applicatio
27	74	4.3	287 12 US-60-079- Sequence 4978, Applic
28	74	4.3	287 15 US-60-089- Sequence 574, Applic
29	73	4.3	383 28 US-08-311- Sequence 8, Applicatio
30	73	4.3	383 31 US-08-446- Sequence 8, Applicatio
31	73	4.3	383 31 US-08-446- Sequence 8, Applicatio
32	73	4.3	383 31 US-08-446- Sequence 8, Applicatio
33	73	4.3	383 31 US-08-451- Sequence 8, Applicatio
34	73	4.3	383 31 US-08-446- Sequence 8, Applicatio
35	66	3.9	339 34 US-08-617- Sequence 9, Applicatio
36	65	3.8	340 34 US-08-617- Sequence 9, Applicatio
37	65	3.8	477 20 US-07-904- Sequence 10, Applicati
38	65	3.8	477 20 US-07-904- Sequence 4, Applicatio
39	62	3.6	340 34 US-08-617- Sequence 9, Applicatio
40	62	3.6	477 20 US-07-904- Sequence 4, Applicatio
41	61	3.6	477 20 US-07-904- Sequence 10, Applicati
42	60	3.5	339 34 US-08-617- Sequence 8, Applicatio
43	52	3.1	259 15 US-60-089- Sequence 504, Applicat
44	52	3.1	392 48 US-09-045- Sequence 2, Applicatio
45	53	3.1	392 48 US-09-040- Sequence 2, Applicatio

RESULT 1

ID US-08-934-254-26 STANDARD; DNA; UNC; 1702 BP.

AC xxxxxx

DT Sequence 26, Application US/08934254

DE Sequence 26, Application US/08934254

CC GENERAL INFORMATION:

CC APPLICANT: Thomas, Terry L.

CC TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A

CC TITLE OF INVENTION: DELTA 6-DESATURASE

CC NUMBER OF SEQUENCES: 27

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Scully, Scott, Murphy & Presser

CC STREET: 400 Garden City Plaza

CC CITY: Garden City

CC STATE: New York

CC COUNTRY: United States

CC ZIP: 11530

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent in Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/934,254

CC FILING DATE:

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Presser, Leopold

CC REGISTRATION NUMBER: 19,827

CC REFERENCE/DOCKET NUMBER: 83832YXWU

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (516) 742-4343

CC TELEFAX: (516) 742-4366

CC TELEX: 230 901 SANS UR

CC INFORMATION FOR SEQ ID NO: 26:

CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1702 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: both
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: DNA (genomic)
 CC FEATURE:
 CC NAME/KEY: CDS
 CC LOCATION: 48..1406
 CC NAME/KEY: CDS
 CC LOCATION: 48..1406
 CC SEQUENCE 1702 BP; 358 A; 471 C; 446 G; 427 T; 0 OTHER.

Query Match 100.0%; Score 1702; DB 43; Length 1702;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 1702; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCCAAAATTTCTTCTCCATCTGGACACAGATCCACAAATGGAGGCGGAAG 60
 1 CCCCAAAATTTCTTCTCCATCTGGACACAGATCCACAAATGGAGGCGGAAG 60
 61 CTAAGAAGTATATCACGCGGAGGAGCTCCGCGCCACAAAGTCCGCGCATCTCTGGA 120
 61 CTAAGAAGTATATCACGCGGAGGAGCTCCGCGCCACAAAGTCCGCGCATCTCTGGA 120
 121 TCTCCATCAGGGAAGGTCTACGACTGCTCTGGTGGGCGGAGGAGCACCAGGCGGCG 180
 121 TCTCCATCAGGGAAGGTCTACGACTGCTCTGGTGGGCGGAGGAGCACCAGGCGGCG 180
 181 AGGTCCCGCTCCTCAGTCTGCGCGGAGGAGCTACGAGCGCTTCAATGGGTACCACC 240
 181 AGGTCCCGCTCCTCAGTCTGCGCGGAGGAGCTACGAGCGCTTCAATGGGTACCACC 240
 241 CGGCGACGCGCTGGCGGATCTGGATCGCTCTTCAACCGGCTACTACTACTCAAGGACT 300
 241 CGGCGACGCGCTGGCGGATCTGGATCGCTCTTCAACCGGCTACTACTACTCAAGGACT 300
 301 TCGAAGTGTGCGGAGATCTCAAGGAGTACCGGAGGCTTTTGAACGAGATGTGCGGCTCG 360
 301 TCGAAGTGTGCGGAGATCTCAAGGAGTACCGGAGGCTTTTGAACGAGATGTGCGGCTCG 360
 361 GGATCTTCGAGAAGGGCCACACATCATGTGACGCTTGTGCGGCTTGGGTCATGA 420
 361 GGATCTTCGAGAAGGGCCACACATCATGTGACGCTTGTGCGGCTTGGGTCATGA 420
 421 TGGCGGCAATCGTCTACGCGCTGTGCGGCTGCGGAGTCCGCGGAGTTCACATGCTCTCG 480
 421 TGGCGGCAATCGTCTACGCGCTGTGCGGCTGCGGAGTCCGCGGAGTTCACATGCTCTCG 480
 481 GCGCACTGCTGGGCTTGTCTGGATCCAGCGCGCTATGTGGGCGCATGACTCCGCGCAT 540
 481 GCGCACTGCTGGGCTTGTCTGGATCCAGCGCGCTATGTGGGCGCATGACTCCGCGCAT 540
 541 ACCAGGTGATGCCAACCGCTGGATACACAGAAATCAGCAACTCATAGCAGCAACATCC 600
 541 ACCAGGTGATGCCAACCGCTGGATACACAGAAATCAGCAACTCATAGCAGCAACATCC 600
 601 TAACCGGAATCAGCATCGCTGGTGGAGTGGACCCACAGCCCAACACCTCGCGCTGCA 660
 601 TAACCGGAATCAGCATCGCTGGTGGAGTGGACCCACAGCCCAACACCTCGCGCTGCA 660
 661 ACAGCTCGACTACGACCCGACCTCCAGCAGATCCCGGATTCGCGCTCTCCACCCGAC 720
 661 ACAGCTCGACTACGACCCGACCTCCAGCAGATCCCGGATTCGCGCTCTCCACCCGAC 720
 721 TCTTCAATCCATCAGCTCGCTTCTTATGGCCGAGTCTTGAATTCGACGAAGTGGAC 780
 721 TCTTCAATCCATCAGCTCGCTTCTTATGGCCGAGTCTTGAATTCGACGAAGTGGAC 780
 781 GGTTCCTAGTACGATACGACACTGGACCTACTACCGCGGTATGATCTTCGCGCGAGTCA 840
 781 GGTTCCTAGTACGATACGACACTGGACCTACTACCGCGGTATGATCTTCGCGCGAGTCA 840

Db 841 ACCTCTTCATCCAGACCTTTTATTCTCCTCCTCACCAGGCGGACGCTCCCTGACGGGCTC 900
 Qy 841 ACCTCTTCATCCAGACCTTTTATTCTCCTCCTCACCAGGCGGACGCTCCCTGACGGGCTC 900
 Db 901 TAAACTTAATGGGTATCGCGGTTTCTGGACGTTGGTCCCGCTCTTCTGATCTTGTCTCC 960
 Qy 901 TAAACTTAATGGGTATCGCGGTTTCTGGACGTTGGTCCCGCTCTTCTGATCTTGTCTCC 960
 Db 961 CGAATGGGCTGAACGGTTCGGTTCCTCATCAGCTTTGCGGTACAGGCGATCCAGC 1020
 Qy 961 CGAATGGGCTGAACGGTTCGGTTCCTCATCAGCTTTGCGGTACAGGCGATCCAGC 1020
 Db 1021 AGTCCAGTTCAGCTCAACCACTTCTCCGCGGACACATACGTTGGGCCCCCCCCAAGGGCG 1080
 Qy 1021 AGTCCAGTTCAGCTCAACCACTTCTCCGCGGACACATACGTTGGGCCCCCCCCAAGGGCG 1080
 Db 1081 ACAACTGGTTCAGAAAGCAGACGAAAGGACGATCCATATCAGTCCGCCACCTGGATGG 1140
 Qy 1081 ACAACTGGTTCAGAAAGCAGACGAAAGGACGATCCATATCAGTCCGCCACCTGGATGG 1140
 Db 1141 ACTGTTCTTTGGTGGGCTGACGTTCCAGTTGGAGCACCACCTTGTTCCTAGGCTGCGCG 1200
 Qy 1141 ACTGTTCTTTGGTGGGCTGACGTTCCAGTTGGAGCACCACCTTGTTCCTAGGCTGCGCG 1200
 Db 1201 GTGGCAGCTTAGGAAGATTGCGCCCTTGGCTCGGAGCTTGTGTAAGACGACGGATGC 1260
 Qy 1201 GTGGCAGCTTAGGAAGATTGCGCCCTTGGCTCGGAGCTTGTGTAAGACGACGGATGC 1260
 Db 1261 CGTATAGGAGCTTCGGGTTTGGGAGCAGCTAAATCTCAGGACAAATCGGACCTGAGGG 1320
 Qy 1261 CGTATAGGAGCTTCGGGTTTGGGAGCAGCTAAATCTCAGGACAAATCGGACCTGAGGG 1320
 Db 1321 ATCGCGGCTTCAGCGCGCTGACCTTAATTCGCGCCCGTCCCTAAGAACTTGGGTATG 1380
 Qy 1321 ATCGCGGCTTCAGCGCGCTGACCTTAATTCGCGCCCGTCCCTAAGAACTTGGGTATG 1380
 Db 1381 GGAAGCTTAAACACCCATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1440
 Qy 1381 GGAAGCTTAAACACCCATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1440
 Db 1441 TATTATGTTATTCACCAATATGAACTGAATGAATGAATGAATGAATGAATGAATGAAT 1500
 Qy 1441 TATTATGTTATTCACCAATATGAACTGAATGAATGAATGAATGAATGAATGAATGAAT 1500
 Db 1501 TTAACCTTCTTAACCTTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 Qy 1501 TTAACCTTCTTAACCTTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 Db 1561 TTATTTCTTATCCCATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1620
 Qy 1561 TTATTTCTTATCCCATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1620
 Db 1621 TTATTTGTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1680
 Qy 1621 TTATTTGTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1680
 Db 1681 TTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1702
 Qy 1681 TTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1702

RESULT 2
 ID US-08-931-575-1 STANDARD; DNA; UNC; 1684 BP.
 AC xxxxxx
 DT

Sequence 1, Application US/08831575
 Sequence 1, Application US/08831575
 CC GENERAL INFORMATION:
 CC APPLICANT: Thomas, Terry L.
 CC APPLICANT: Li, Zhongsen
 CC TITLE OF INVENTION: AN OLEOSIN 5' REGULATORY REGION FOR THE
 CC TITLE OF INVENTION: MODIFICATION OF PLANT SEED LIPID COMPOSITION

CC REGISTRATION NUMBER: 19,827
CC REFERENCE/DOCKET NUMBER: 83832YXWVU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (516) 742-4343
CC TELEFAX: (516) 742-4366
CC TELEX: 230 901 SANS UR
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1685 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: both
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 1685 BP; 431 A; 277 C; 357 G; 620 T; 0 OTHER.

Query Match 10.1%; Score 172; DB 43; Length 1685;
Best Local Similarity 57.7%; Pred. No. 4.24e-178;
Matches 697; Conservative 0; Mismatches 507; Indels 3; Gaps 2;

42 CAATGGCTGCTCAAAATCAAGAAATACATTTACCTCAGATGAACCTCAAGAACACACATTAAC 101
46 CAATGGAGGCGGAGCTAAGAAATATATACGCGGAGGAGCCTCCGCGGCCACACAAGT 105
102 CCGGAGATCTATGATCTCGATTCAAGGAAAGCCTATGATTTTCGGATTGGGTGAAG 161
106 CCGCGATCTCTGGAATCCATCCAGGCAAGGTCTACGACTGCTCTCGGTGGGCGCGG 165
162 ACCATCCAGGTGCGAGCTTCCCTTTGAAGAGTCTTGTGTCAGAGAGTCACTGATGCAT 221
166 AGCACCCGCGGCGAGGTCCTCCCTCCTCAGCTGCGCGGCCAGGAGCTCACCGACGCT 225
222 TTGTTGATTCATCTCGCTCTACATGGAAGATCTGTATGAATTTTACTGGGTATT 281
226 TCATTTGGTACCAACCGGCGAGGCTGCGGAGTCTGATGATCCGCTCTTACCGGCTACT 285
282 A-TT--CTTAAGATTACTCTGTTCTCAGGTTTCTAAGATTATAGAACTTAGAGCTGTTG 338
286 ACTACCTCAAGGACTTCGAAGTTCGAGAGATCTCCAAGGACTACCGGAGGCTTTGAACG 345
339 AGTTTCTTAAATGGGTTTGTATGACAAAAAGGTCTATATTTTGCACACTTTGTGCT 398
346 AGATGTCGCGGTCGGGATCTTCAGAGAGAGGCCACCATCATGTGACGTTCTGTCG 405
399 TTAAGCAATGCTGTTTGTATGATGATGTTTATGGGGTTTGTGAGGGTGTGTTTG 458
406 CGGTTGCGGTCATGATGCGGCAATCTCTACGCGCTGCTGCGCTCGGAGTCCGCTCGAG 465
459 TACATTTGTTTCTGGGTGTTTGTAGGGGTTTCTTGGATTCAGATGTTGGATTGGAC 518
466 TTCATGCTCTGCGGCGCACTGCTGGGCTTGTGTGATCCAAAGCGCGGTATGTTGGGCC 525
519 ATGATGTCGGCATTAATATGATGCTGCTGATTCAAGGCTTAATAAGTTTATGGTATT 578
526 ATGACTCGGCCATACAGGTGATGCCACCCCTGATACAAAGAAATACGCAACTCA 585
579 TTGCTCAAAATGCTTTTTCAGGAATAAGTATTGTTGGTGAATTAAGAAATCAATATGAC 638
586 TAGCAGCAACATCTAACCAGATCAGATCGGCTGGTGAAGTGAACCCACACACGCC 645
639 ATCATTGCTGTAATGAGCTTGAATATGACCTGATTTTACAATATATACCATTCCTTG 698
646 ACCACCTCGCTGCAACAGGCTGACTACGACCCCGACCTCCAGCATCCCGCTATTTCG 705
699 TTGTTCTCTCCAGTTTGTGTTTCACTCACTCTCTTCTATGATAAGAAAGTTGACTT 758
706 CCGTCTCCACCGGACTCTTCAACTCCATCACTCGGCTCTCTATGGCGGAGTCTGAAAT 765
759 TTGACTCTTTTCAAGATCTTTTGAAGTTATCAACATTTGACATTTTACCTATTATGT 818
766 TCGACGAAGTGGACGCTTCTTAGTACGCTACGAGCTGAGCACTACTACCCGCTCATGA 825
819 GTGCTGCTAGGCTCAATATGATGATGATCAATCTCTCATAATGTTTGTGACCAAGAAATG 878

QY 826 TCTTCGGCGAGTCAACCTCTTCTCATCCAGACCTTTTATTGCTCTCTCACCAGCGCGAGC 885
Db 879 TGTCTATTCAGACTCAGAACTCTTGGATGCTCCTAGTCTTCTCGATTTGGTACCGTTGC 938
QY 886 TCCCTGACCGGCTCTAAACTTAATGGGTATCGGGTTTCTGGACGTGGTTCCCGCTCT 945
Db 939 TTGTTTCTTGTTCCTTAATTTGGGTGAAAGAATATGTTTGTATTTCAGATTATCAG 998
QY 946 TCGTATCTTGTCTCCGAACCTGGCTTGAACGGTTTCGGGTTCTCTCATCAGCTTTGCGG 1005
Db 999 TGAAGTGAATCAACAAAGTTTCAGTTCTCTTGAACCACTTCTCTTCAAGTTTATGTTG 1058
QY 1006 TCAGGGCGATCCAGCAGCTCCAGTTCCAGTCAACCACTTCTCCGGCCACACATACGTTG 1065
Db 1059 GAAAGCCTAAAGGAATAATTGGTTTGAGAAACAACAGATGGGACACTTGACATTTCTT 1118
QY 1066 GCCCCCCCAAGGGGCAACCTGTTTCGAGAGAGCAGAGAAAGGAGCATCGATATCAGT 1125
Db 1119 GTCCTCTTGTGATGATGTTTTCATGTTGGTGGATTTGCAATTCGAATTCAGATCATTTGT 1178
QY 1126 GCCCAGCTGGATGAGTGGTTCTTGTGGCTGCACTCCAGTTGGAGCACCACCTTGT 1185
Db 1179 TTCCCAAGATGCTAGATGCAACCTTAGGAAATCTCGCCCTAGTATCGAGTTATGCA 1238
QY 1186 TCCCTAGCTGCCGCGTGGCAGCTTAGGAAGATTGGCCCTTGGCTCGGAGCTTGTGTA 1245
Db 1239 AGAAACA 1245
QY 1246 AGAGCA 1252

RESULT 5
ID US-60-086-722-1966 STANDARD; DNA; UNC; 291 BP.
AC xxxxxx
DT
DE Sequence 1966, Application US/60086722
CC Sequence 1966, Application US/60086722
CC GENERAL INFORMATION:
CC APPLICANT: Lalgudi, Raghunath
CC APPLICANT: Ito, Laura Y.
CC APPLICANT: Sherman, Bradley
CC TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
CC TITLE OF INVENTION: CORN EARS
CC NUMBER OF SEQUENCES: 7600
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 8.0 for Windows/MS-DOS 6.2
CC CURRENT APPLICATION DATA: US/60/086,722
CC APPLICATION NUMBER: US/60/086,722
CC FILING DATE: HEREWITH
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CERRONE, MICHAEL C.
CC REGISTRATION NUMBER: 39,132
CC REFERENCE/DOCKET NUMBER: PL-0017 P
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 855-0555
CC TELEFAX: (650) 845-4166
CC INFORMATION FOR SEQ ID NO: 1966:
CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 291 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: 700551823H1
SQ SEQUENCE 291 BP; 47 A; 88 C; 80 G; 75 T; 1 OTHER.

Query Match 6.8%; Score 116; DB 14; Length 291;
Best Local Similarity 70.8%; Pred. No. 2.32e-106;
Matches 204; Conservative 0; Mismatches 83; Indels 1; Gaps 1;
Db 2 CCCGTGTGTTGGCTCCCTCCCAATTTGGTGGAGAGGTCGCGTTTGTCTTTTCAG 61
QY 938 CCCGCTCTCTGATCTGTCTCCCAATTTGGTGGAGAGGTCGCGTTTGTCTTCATCAG 997
Db 62 CTTACCACTCTCGGGATTCAGCAGCTCAATTTCTGCTGAACCACTTCTCTCGCGAGGT 121
QY 998 CTTTCGGGTTCACGGGATTCACAGCTCCAGTTCAGCTCAACCACTTCTCGGGGACAC 1057
Db 122 GTATCTCGGGCCACCCAGGGCAATCTGCTTTGAGAGCAGCGGACGACGCTCGA 181
QY 1058 ATACGTGGGCCCCCAGGGGCAACTGTTTCGAGAAGCAGACGAAAGGACGATCGA 1117
Db 182 CATCTGTGCTCTCTTGGATGGATGTTCCACGTCGCTGCGCTTCCAGATTGAGCA 241
QY 1118 TATCAGTGGCCACCGTGGATGGTCTTGTGGGCTGCACTTCCAGTTGGAGCA 1177
Db 242 CCATCTGTTCCCNCTACCTAGTGGCCACCTCCGTAAG-TTGGGCC 288
QY 1178 CCACCTGTTCCCTAGCTGCCCGTGGGAGCTTAGGAAGATTGGCC 1225

RESULT 6

ID US-60-086-856-2042 STANDARD; DNA; UNC; 276 BP.
AC xxxxxx

DE Sequence 2042, Application US/60086856
CC Sequence 2042, Application US/60086856
CC GENERAL INFORMATION:
CC APPLICANT: Lalgudi, Raghunath
CC APPLICANT: Ito, Laura Y.
CC APPLICANT: Sherman, Bradley
CC TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
CC TITLE OF INVENTION: CORN EMBRYO
CC NUMBER OF SEQUENCES: 8690
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 8.0 for Windows/MS-DOS 6.2
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/086,856
CC FILING DATE: HEREWITH
CC CLASSIFICATION:
CC PRIOR APPLICATION NUMBER:
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CERRONE, MICHAEL C.
CC REGISTRATION NUMBER: 39,132
CC REFERENCE/DOCKET NUMBER: PL-0016 P
CC TELEPHONE: (650) 855-0555

CC TELEFAX: (650) 845-4166
CC INFORMATION FOR SEQ ID NO: 2042:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 276 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: 700261349H1
SQ SEQUENCE 276 BP; 49 A; 85 C; 74 G; 67 T; 1 OTHER.

Query Match 6.5%; Score 110; DB 14; Length 276;
Best Local Similarity 72.4%; Pred. No. 7.87e-99;
Matches 176; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
Db 3 GATTCAGCAGCTCCCAATTTGCTGCTGAACCACTTCTGCTCCGACGTGTATGTCGGGCCACC 62
QY 1013 GATTCAGCAGCTCCCAATTTGCTGCTGAACCACTTCTGCTCCGACGTGTATGTCGGGCCACC 1072
Db 63 CAAGGCGAATGACTGGTTTGAAGACAGCAGCGACGCTCGACATCCTGTGCTCTCC 122
QY 1073 CAAGGCGAATGACTGGTTTGAAGACAGCAGCGACGCTCGACATCCTGTGCTCTCC 1132
Db 123 TTGGATGGATTTCCACGGTGGCTGCAGTTCAGATTGAGCACCATCTGTTTCCCG 182
QY 1133 GTGGATGGATTTCCACGGTGGCTGCAGTTCAGATTGAGCACCATCTGTTTCCCTAG 1192
Db 183 CCTACTAGTTCGCCACCTCCGTAAGTTGCGCTGCCGTCGCGAGCTTTGCAAGAAGCA 242
QY 1193 GTCGCGGTGGGAGCTTAGGAAGATTGCGCCCTTGGCTCGGACTTGTGTAAAGAAGCA 1252
Db 243 TGG 245
QY 1253 CGG 1255

RESULT 7

ID US-60-089-804-1341 STANDARD; DNA; UNC; 267 BP.
AC xxxxxx

DE Sequence 1341, Application US/60089804
CC Sequence 1341, Application US/60089804
CC GENERAL INFORMATION:
CC APPLICANT: Stuve, Laura L.
CC APPLICANT: Gooding, Douglas H.
CC APPLICANT: Ito, Laura Y.
CC APPLICANT: Sherman, Bradley K.
CC TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED
CC TITLE OF INVENTION: FROM CORN EMBRYO
CC NUMBER OF SEQUENCES: 2470
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: Word Perfect 6.1/MS-DOS 6.2 software
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/089,804
CC FILING DATE: HEREWITH
CC CLASSIFICATION:
CC PRIOR APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CERRONE, MICHAEL M.
CC REGISTRATION NUMBER: 36,132

CC	REFERENCE/DOCKET NUMBER: PL-0007 P
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (650) 855-0555
CC	TELEFAX: (650) 845-4166
CC	INFORMATION FOR SEQ ID NO: 1738:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 298 base pairs
CC	TYPE: nucleic acid
CC	STRANDEDNESS: single
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: cDNA
CC	IMMEDIATE SOURCE:
CC	CLONE: 700100371H1
SO	SEQUENCE 298 BP: 56 A; 90 C; 82 G; 69 T;

Query Match

Best Local Similarity 71.7%;
Matches 160; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Db 1 CTTCTCGTCCGAGTGATGTGGGCCACCCCAAGGCCAATGACTGGTTTGAGAAGCAGAC 60
||||| |||| | | | | | | | | | | | | | | | | | | | | |
Ov 1043 CTTCTCGGGCGACATAGTGGGGCCCCCCC AAGGGCGACAAGTGGTCGAGAAGCAGAC 1102

db 61 GGCAGGCAGGCTCGACATCCCTGTGCTCTCCTTGGATGGATTGGTTCCACGGTGGCCTGCA 120

09 1103 GAAAGGACGATCGATATCACGTGCCACCCTGGATGGACTGGTTCCTTTGGTGGGTGCA 1162

d_b 121 CTTCCACGATTTCAGNCACCACTCCTTCCCTCCATCCTCTCC 180

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT 9

AC XXXXXX
00 00 000 010 4177 STANDARD, DMS, UNC, 304 BF.

Sequence 4177, Application US/60089810

DE sequence 4177, Application US/60089810

CC GENERAL INFORMATION:
CC APPLICANT: Joseph R. Byrum

CC APPLICANT: Joseph K. Byrum

CC APPLICANT: Michael D. Thompson

33 TITLE OF INVENTION: Nucleic Acid M

33 TITLE OF INVENTION: Associated Wit

CC NUMBER OF SEQUENCES: 5409
CC CORRESPONDENCE ADDRESS.

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Lawrence M. Lavin, Jr.

CC ADDRESSEE: Lawrence M. Lavin, Jr.
CC STREET: BB4F

CC STREET: BB4F
CC STREET: Monsanto Company

STREET: Mousanto Company
STREET: 700 Chesterfield Parkway

STREET: 700 Chesterfield Parkway
CITY: St. Louis

CC
CC
CITY: ST. LOUIS
STATE: MO

CC STATE: MO
CC COUNTRY: United States

CC
CC
COUNTRY: United States
ZIP: 63198

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette (3.5 in, 1.44 MB)

CC
CC
CC
MEDIUM LIFE: DISKette (3.5in, 1
COMPUTER: IBM PC/XT/AT, IBM PS/2

COMPUTER: IBM PC/AT/A1, IBM PS/2
OPERATING SYSTEM: Windows 95/NT

CC
CC
CC OFFERING SYSTEM: WINDOWS 95/NL
SOFTWARE: WinZip 6.3

CC SOFTWARE: WINZIP 6.5
CC CURRENT APPLICATION DATA:

CC
CC
CURRENT INFORMATION DATA:
APPLICATION NUMBER: US/60/089,81

CC FILING DATE:
CC CLASSIFICATION:

CC
CC
CC
CLASSIFICATION.
CLASSIFICATION.
ATTORNEY/AGENT INFORMATION:
ATTORNEY/AGENT INFORMATION:

CC NAME: Lavin Jr., Lawrence M.
CC ATTORNEY/AGENT INFORMATION.
CC

REGISTRATION NUMBER: 30,768

REFERENCE/DOCKET NUMBER: 38-21(1

TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (314) 737-6670

CC
TELEFAX: (314) 737-6047

[illegible]

```

RESULT 12.
ID US-60-086-722-3256 STANDARD; DNA; UNC; 266 BP.
AC
xxxxx
DT
DE Sequence 3256, Application US/60086722
DE Sequence 3256, Application US/60086722
CC
GENERAL INFORMATION:
CC APPLICANT: Lalgudi, Raghunath
CC APPLICANT: Ito, Laura Y.
CC APPLICANT: Sherman, Bradley
CC TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
CC TITLE OF INVENTION: CORN EARS
CC NUMBER OF SEQUENCES: 7600
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 8.0 for Windows/MS-DOS 6.2
CC
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/086,722
CC FILING DATE: HERewith
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CERRONE, MICHAEL C.
CC REGISTRATION NUMBER: 39,132
CC REFERENCE/DOCKET NUMBER: PL-0017 P
CC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 855-0555
CC TELEFAX: (650) 845-4166
CC INFORMATION FOR SEQ ID NO: 3256:
CC
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 266 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC IMMEDIATE SOURCE:
CC CLONE: 700611205H1
CC
CC SEQUENCE 266 BP; 47 A; 81 C; 83 G; 54 T; 1 OTHER.
CC
Query Match 5.3%; Score 90; DB 14; Length 266;
Best Local Similarity 76.5%; Pred. No. 4.52e-74;
Matches 137; Conservative 0; Mismatches 41; Indels 1; Gaps
Db 1 GGTTCAGAAGCAGACGGCAGGACCTCGACATCTCTGTGCCCTCTCTGGATGGATTGGT 60
QY 1087 GGTTCAGAAGCAGACGAAAGGAGCATCATCAGTGCACCGCTGGATGGATGGT 114
Db 61 TCCATGGCGCCTGCAGTTCACAGATTGACACCATCTGTTCCCGCCCTCGCTCGTGC-C 119
QY 1147 TCATTGGTGGCTGCAGTTCAGATTGGAGACACCACTGTTCCCTAGGCTCCCGCGTGGGC 120
Db 120 ACCTTCGGAAGTGTGGCCGCGGCGTGCATCTTTGCAAGAAGCATGGCGCTGACCTAT 178
QY 1207 AGCTTAGGAAGATTGGCCGCTTGGCTCGGGACTTGTGTAAAGAAGCAGCGGATCCGCTAT 1265
RESULT 13
ID US-08-466-194-14 STANDARD; DNA; UNC; 7218 BP.
AC
xxxxx
DT

```

DE Sequence 14, Application US/08466194
CC Sequence 14, Application US/08466194
CC GENERAL INFORMATION:
CC APPLICANT: DORNER, F.
CC APPLICANT: SCHEIFLINGER, F.
CC APPLICANT: FALKNER, F. G.
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, NW, Suite 500
CC CITY: Washington
CC STATE: DC
CC COUNTRY: USA
CC ZIP: 20007
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/466,194
CC FILING DATE: 06-JUN-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/232,463
CC FILING DATE: 22-APR-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/935,313
CC FILING DATE: 26-AUG-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 28-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/201 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC TELEX: 904136
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pTZ9pt-Fls
CC SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match 5.3%; Score 91; DB 31; Length 7218;
Best Local Similarity 0.3%; Pred. No. 2.70e-75;
Matches 1; Conservative 229; Mismatches 139; Indels 0; Gaps 0;

Db 1067 YY 1126
QY 634 CCACACGCCGCCACCCCTCGCTGCAAGCCTGACTACGACCCGACCTCCAGCACA 693
Db 1127 YY 1186
QY 694 TCCCGTATCGCGGCTCCACCGGCTTCAACTACCTACCTCGGCTTCTATGGCC 753
Db 1187 YY 1246
QY 754 GAGTCTGAATCGACGAGTGGCAGGTCTCTAGTACGACGACTGACCTACT 813
Db 1247 YY 1306
QY 814 ACCCGGTGATGCTTCGCGCGAGTCAACTCTCATCCAGACCTTTTATGCTCTCA 873
Db 1307 YY 1366
QY 874 CGAGCGGACGCTCCCTGACCGGCTCTAAACTTAATGGGTATCGCGGTTTCTGACGT 933

Db 1367 YY 1426
QY 934 GGTCCGCTCTTCGTATCTGTCTCCGAACTGCCTGACGCTGCGGTTCGCTCTCA 993
Db 1427 YYYYYYYYG 1435
QY 994 TCAGCTTTG 1002

RESULT 14

ID US-08-569-578-14 STANDARD; DNA; UNC; 578 BP.
AC xxxxxx
DT

Sequence 14, Application US/08569578
Sequence 14, Application US/08569578
GENERAL INFORMATION:
CC APPLICANT: KAVANAUGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC APPLICANT: WILLIAMS PH.D., LEWIS T.
CC TITLE OF INVENTION: IDENTIFICATION, CLONING, AND USE OF
CC TITLE OF INVENTION: SHC-ASSOCIATED INOSITOL POLYPHOSPHATE 5-PHOSPHATASE
CC TITLE OF INVENTION: (SAPTAPE)
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHIRON CORPORATION
CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/569,578
CC FILING DATE: 08-DEC-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-601-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 578 BP; 37 A; 11 C; 62 G; 65 T; 403 OTHER.

Query Match 5.2%; Score 89; DB 33; Length 578;
Best Local Similarity 12.2%; Pred. No. 7.53e-73;
Matches 66; Conservative 259; Mismatches 207; Indels 9; Gaps 9;

Db 25 KTSDDSKVKAHSHRSVTVKASGKMKVDYSGKKSKDGDGSKYSHKKSKNKNVYTKRKY 84
QY 826 TCTTCGCGGAGTCAACCTCTTCATCCAGACCTTTTATGCTCCTCACAGGCGCAGC 885
Db 85 VADSKRGCMKNKHSMTGTWNMGNAKTSWSKGGKTRDSDADYH-DYVGTDSSKWKHS 143
QY 886 TCCTGACCGCGCTCTAAACTTAATGGGTATCGCGGTTTCTGACGCTGCTCCCGC-TC 944
Db 144 VTKVTAHTNVRVAKHNRSCHTDNVTGANTGNKGVSMNGTSGVNSHTSGSKRRNN 203
QY 945 TTGCTATCTGTCTCCGAACTGGCTGCGGTTCGCTCCTCATCAGCTTTGCG 1004
Db 204 YNMRAGDKKSNTHRTHWGDNYRVDTWATKKYADSHDTRRKVHTATYRTRDKYATKKAT 263

Query Match 5.2%; Score 89; DB 34; Length 578;
Best Local Similarity 12.2%; Pred. No. 7.53e-73;
Matches 66; Conservative 259; Mismatches 207; Indels 9; Gaps 9;

Db 25 KSTSSDKVKAHSHRSYVTVKASGKMKVDVSGKSKSDGSDKYSKHKSKSNKVVTKKRY 84
QY TCTTCGGCGGAGTCAACCTCTTCACAGACCTTTTATGCTCCTCACCAGCGCGAG 885
Db 85 VADSKKRCMKNKHSDMTGTWNMGNAKTSKSGKTRDDSDADYH-DYVGTDSKWKHSTS 143
QY TCCCTGACCGCGCTCTAAACTTAATGGGTATCGCGGTTTCTGGACGTGTTCCCGC-TC 944
Db 144 VTKIVAHVWNRVVAKHNRSHCHTDNKTGANTGNKAGVSMNGTSGVNSHTSGSKRRNN 203
QY TCCCTGATCTTGTCTCCGAACTGGCTGAACGGTTTCGGGTTGCTCCTCATCAGCTTTGG 1004
Db 204 YNNRAGDKSNTHRTHWGDNYRVDTWATKKVADSHDTRRKVHTATYRTRDKYATKKAT 263
QY 1005 GTCACGGGATCCACAGCTCCAGTTCAGCTCAACCACTTCTCCGGGACACATACGTG 1064
Db 264 GMKYNWCDRVWKSYYHVVCYSGTSDMTSDHSVATAGVTSVSKNGTVDSGRCYATKTK 323
QY 1065 GSCCCCCCAAGGGCGACACTGGTTCGAGAAGCAGACGAAGGACGATCGA-TATCAC 1123
Db 324 STKYHSSCVKSGNSGVVVKFKSDYDHSKSSDSYGGCAR-ATTYTH-HGTGCHKT 381
QY 1124 GTGCCACCGTGGATGGAGTGGTTTGGTGGGCTGAGTTCCAGTGGAGCACCACTT 1183
Db 382 SGKTRKYDVKTRDSSG-KTKSTSHDMKWTSRACSGSSTNNYMGVGMHVKTSDTAWSYD 440
QY 1184 GTCCCTAGCTGCGGCGGACCTTAGGAAGATTGCGCCCTTGGCTCGGAGCTTTGTG 1243
Db 441 KDSGCRGSGTSGKSTANRGRTSRSDGKNAGDTRKMYGSSSKARKDSKMRKCGSSVTKA 500
QY 1244 TAAGAAGCAGCGGATCCGCTATAGAGCTTCGGGTTTGGGACGACGCTAAT-GTCAGGA 1302
Db 501 ADGGGKVARRSTCSSAGRAAGDKSGKKTVSSAVAKRKSRSNTTTRVKSVAHVHSGKRD 560
QY 1303 CAATTGGAGCGCT-GAG-GGATGCGCGGTTTCAGGCGGTGACCTTAATTCGGCCCCGTG 1360
Db 561 Y 561
QY 1361 C 1361

RESULT 15

ID US-08-624-190-14 STANDARD; DNA; UNC; 578 BP.

AC xxxxxx

DT

Sequence 14, Application US/08624190

CC Sequence 14, Application US/08624190

CC GENERAL INFORMATION:

CC APPLICANT: KAVANAUGH PH.D., MICHAEL

CC APPLICANT: POT PH.D., DAVID

CC APPLICANT: WILLIAMS PH.D., LEWIS T.

CC TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE

CC TITLE OF INVENTION: 5-PHOSPHATASE

CC NUMBER OF SEQUENCES: 38

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CHIRON CORPORATION

CC STREET: 4560 HORTON STREET

CC CITY: EMERYVILLE

CC STATE: CA USA

CC COUNTRY: USA

CC ZIP: 94608

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/624,190

CC FILING DATE:

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: GREEN, GRANT

CC REGISTRATION NUMBER: 31,259

CC REFERENCE/DOCKET NUMBER: 1182.003

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 510-601-2706

CC TELEFAX: 510-655-3542

CC INFORMATION FOR SEQ ID NO: 14:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 578 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

SQ SEQUENCE 578 BP; 37 A; 11 C; 62 G; 65 T; 403 OTHER.

Search completed: Thu Feb 18 18:55:18 1999
Job time : 3658 secs.

W P E R E H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 16:22:44 1999; MasPar time 2188.90 Seconds
1392.438 Million cell updates/sec
Linear output not generated.

Title: >US-08-934-254-26
Description: (1-1702) from US08934254.seq
Perfect Score: 1702
N.A. Sequence: 1 CCCCAAAATTTTCATTGTTTTTTGGTAAAAA 1702
Comp: GGGGTTTTTAAAGTACAAAAAAACCATTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est56
Database: genbank-est109

5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21
18:gb_est22 19:gb_est23 20:gb_est24 21:gb_est25 22:gb_est26
23:gb_est27 24:gb_est28 25:gb_est29 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 11.586; Variance 2.268; scale 5.110

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
1	172	10.1	556 11	H36955 15084 Lambda-PRL2 Arab	3.96e-283
2	156	9.2	408 11	W43255 22634 Lambda-PRL2 Arab	4.38e-250
3	142	8.3	480 11	T43806 6069 Lambda-PRL2 Arabi	2.18e-221
4	127	7.5	455 11	T44542 7805 Lambda-PRL2 Arabi	6.45e-191
5	88	5.2	318 20	ATTS4723 A. thaliana transcribe	1.28e-113
6	64	3.8	562 11	N37558 18785 Lambda-PRL2 Arab	2.21e-68
7	60	3.5	252 12	AA754459 97SN1787 Rice Immature	3.94e-61
8	58	3.4	252 12	AA754459 97SN1787 Rice Immature	1.52e-57
9	55	3.2	247 12	AA754458 97SN1784 Rice Immature	3.18e-52
10	51	3.0	352 11	T42569 5832 Lambda-PRL2 Arabi	3.01e-45
11	50	2.9	597 16	T15209 crs698 Ricinus communi	1.59e-43
12	48	2.8	247 12	AA754458 97SN1784 Rice Immature	4.09e-40
13	46	2.7	2275 11	AF034173 Homo sapiens ntcon2 co	9.56e-37

LOCUS	DEFINITION	1	H36955	556 bp	mrna	EST	30-DEC-1997
14	34	2.0	448 17	AI107749	GH05586.5prime GH Dros	1.05e-17	
15	34	2.0	613 17	AI107703	GH05526.5prime GH Dros	1.05e-17	
16	32	1.9	470 9	AA527400	ng37b07.s1 NCI_CGAP-Co	8.75e-15	
17	33	1.9	2275 11	AF034173	Homo sapiens ntcon2 co	1.12e-16	
18	30	1.8	505 7	C25669	Dictyostellium discoide	5.81e-12	
19	30	1.8	611 12	AA736027	GM10804.5prime GM Dros	5.81e-12	
20	30	1.8	616 12	AA735895	GM10052.5prime GM Dros	5.81e-12	
21	30	1.8	698 14	AA941503	LD25669.5prime LD Dros	5.81e-12	
22	29	1.7	310 8	AA279374	zs84e05.s1 NCI_CGAP-GC	1.37e-10	
23	29	1.7	370 12	AA767321	nz65f07.s1 NCI_CGAP-GC	1.37e-10	
24	29	1.7	564 18	AI135916	GH13838.5prime GH Dros	1.37e-10	
25	29	1.7	591 17	AI109930	GH09329.5prime GH Dros	1.37e-10	
26	29	1.7	833 20	W28140	44e3 Human retina cDNA	1.37e-10	
27	27	1.6	100 24	AA296594	EST11122 Uterus Homo s	6.14e-08	
28	28	1.6	300 9	C35646	C.elegans cDNA clone y	3.00e-09	
29	28	1.6	366 28	AQ133540	HS_3051.A1.D08.MR.CIT	3.00e-09	
30	28	1.6	378 13	AA804782	Ob98h06.s1 NCI_CGAP-GC	3.00e-09	
31	27	1.6	398 8	AA279631	zs86h09.r1 NCI_CGAP-GC	6.14e-08	
32	27	1.6	420 28	AQ210456	HS_2202.B2.A08.MF.CIT	6.14e-08	
33	28	1.6	429 20	W53753	md12f10.r1 Soares mous	3.00e-09	
34	27	1.6	452 9	AA531459	nj63c11.s1 NCI_CGAP-Pr	6.14e-08	
35	27	1.6	528 10	AA642379	ns25e09.s1 NCI_CGAP-GC	6.14e-08	
36	27	1.6	547 8	AA262103	zs19a05.s1 NCI_CGAP-GC	6.14e-08	
37	27	1.6	565 12	C84893	Dictyostellium discoide	6.14e-08	
38	28	1.6	681 22	AA148344	JM00M066.QM3 Miracidia	3.00e-09	
39	28	1.6	719 26	B21675	F17P15-Sp6 IGF Arabidop	3.00e-09	
40	27	1.6	752 26	B21679	F8L9-Sp6 IGF Arabidop	6.14e-08	
41	27	1.6	802 18	AU035600	Mus musculus cDNA, clo	6.14e-08	
42	28	1.6	809 26	B10155	F4F14-Sp6.1 IGF Arabid	3.00e-09	
43	27	1.6	858 26	B18671	T29F5-Sp6 TAMU Arabidop	6.14e-08	
44	27	1.6	878 26	B20048	T16G2-T7 TAMU Arabidop	6.14e-08	
45	28	1.6	900 26	B08164	F11011-T7 IGF Arabidop	3.00e-09	

ALIGNMENTS

RESULT 1
LOCUS 15084 Lambda-PRL2 Arabidopsis thaliana cDNA clone 181D1217, mRNA
DEFINITION H36955 sequence.
ACCESSION H36955
NID 9906454
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta;
Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 556)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@bm.cl.msu.edu
Seq primer: T7 dye primer.
FEATURES
Location/Qualifiers
1..556
/organism="Arabidopsis thaliana"
/strain="var columbia"
/note="Vector: lambda Zip-Lox; Site.1: Sal; Site.2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7

sequence.
 ACCESSION F13728
 NID 975104
 KEYWORDS EST; expressed sequence tag; partial cDNA sequence.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
 Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
 Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 318)
 Morris,P.C., Guerrier,D., Barbet,J.C. and Giraudat,J.
 Direct Submission
 Submitted (24-MAR-1995) CNRS, GDR-1003 ACS, INRA, Laboratoire de
 Biologie Moleculaire, BP 27, 31326 Castanet-Tolosan cedex, France.
 E-mail: gdr-sv@toulouse.inra.fr. On behalf of: CNRS, ISV-CNRS, BP 1
 Bat 23 91198 Gif/Yvette, France.
 E-mail: giraudat@refle.isv.cnrs-gif.fr
 2 (bases 1 to 318)
 CNRS.
 The Arabidopsis thaliana transcribed genome: the GDR cDNA program
 Unpublished
 Cloning vector: Lambda ZAPII non-oriented;
 Physiological conditions: greenhouse plants.
 full automatic.
 FEATURES
 source
 1. 318
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="YAY347"
 /tissue_type="Green siliques of A. thaliana ecotype
 columbia"
 /clone_lib="Gif-Seed4"
 68 a 100 c 57 g 89 t 4 others
 BASE COUNT
 ORIGIN
 Query Match 5.2%; Score 88; DB 20; Length 318;
 Best Local Similarity 64.3%; Pred. No. 1.28e-113;
 Matches 189; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
 Db 25 GGTGCGCCGATGTCCTCGGGAATCTCTACGGTGTGTTGGCTTCCTCCGTCGCTNCCC 84
 QY 407 CGTTCGGTATGATGGCGGAATCGTACGGCGTGTGGCTGCGGATCGTGGAGT 466
 Db 85 TCACCAATATGCGCGCGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 144
 QY 467 TCACATGCTCTGCGCGCACTGCTGGGCTGTGCTGGATCCAAAGCGGTATGTGGGCA 526
 145 CGATTCTGTCATTACGTTATCATGTCGAACAAATCTAATACANATTCGCTCAGCTTCT 204
 527 TGACTCCGCCATTCACAGGTATGCCAACCCGCTGGATACAAAGATACCGCAACTCAT 586
 Db 205 CTCGGTAACCTGTCACCGGAATCTAAATCGCTGGTGGAAATGACATCAAAATGCTCA 264
 QY 587 AGCAGCAACATCTTACCGGAATCAGCATCGCTGGTGGAGTGGAGTGGAGTGGAGT 646
 Db 265 TCATCTAGCTGTACAGCTCGATTCAGGTCCTCAATCTCAACATCCCTCGT 318
 QY 647 CCACCTCGCTGCAACAGCTCGATAGGACCCCGACCTCCAGCAGCATCCCGGT 700
 RESULT 6
 LOCUS N37558 562 bp mRNA EST 05-JAN-1998
 DEFINITION 18785 Lambda-PRL2 Arabidopsis thaliana cDNA clone 205087, mRNA
 sequence.
 ACCESSION N37558
 NID 91158700
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryotes; Viridiplantae; Charophyta/Embryophyta group;
 Embryophyta; Tracheophyta; seed plants; Magnoliophyta;
 Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 562)
 Newmann,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
 McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,
 Retzel,E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@lms.cl.msu.edu
 Seq primer: T7 dye primer.
 Location/Qualifiers
 1. 562
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRL's lambda Zip-Lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dT primed cDNA."
 /db_xref="taxon:3702"
 /clone="205087"
 /clone_lib="Lambda-PRL2"
 116 a 141 c 128 g 156 t 21 others
 BASE COUNT
 ORIGIN
 Query Match 3.8%; Score 64; DB 11; Length 562;
 Best Local Similarity 57.3%; Pred. No. 2.21e-68;
 Matches 270; Conservative 0; Mismatches 197; Indels 4; Gaps 3;
 Db 29 CATCCCGGAGGACACGGTATCTCAATCTGTTGGTCAAGACGTACCGATGCTTTC 88
 QY 168 CACCCCGGGGAGGTCCGCTCTCTAGTCTGGCGGCGGAGGACGTACCGACGCTTC 227
 Db 89 ATCGCATTTATCCCGGAACCGCTTGGACCATCTCGACCATCTCTTCCACGGTTACCA 148
 QY 228 ATTGGGTACCAACCGGCGGCGGTGGCGCATCTGGATCCGCTCTTCCACGGCTACT 287
 Db 149 -A--TCAGAGATTTCCAGTCTCCGAAGTCTCACGGATACCGTGTATGGTGCAG 205
 QY 288 TACCTCAAGGACTTCGAAGTGTGGAGATCTCCAAGGACTACGGAGGCTTTTGAACG 347
 Db 206 TTTCGTAACCTCGGTCTCTTCCAAACAAAGTCAAGTCTCTTCTACACTTAGCTTTC 265
 QY 348 ATGTCCGGTCCGGATCTTCGAGAAAGAGGCGCCACCATCATGTGGAGCTTCGTCG 407
 Db 266 GTCGGCGCATGTTCTCTNGGAGTCTCTACGGTGTGTTGGCTTGTACCTCCGCTTCG 325
 QY 408 GTTGGGTATGATGGCGCATCTCTACGGCTGTCTGGCGTGTGGAGTCCGTCGGAGT 467
 Db 326 AACCAAAATCGCGCGGCTTCTTCGGGTCTCTCTGGATCCAGAGNGNTATACAGGCA 385
 QY 468 CACATGCTCTGGCGCACTGCTGGCT-TGCTGTGGATCCAAAGCGCGTATGTGGGCA 526
 Db 386 CGTNTTGGTATACGGTATCATGTNGACAATCTNTTAAACAGTTTCGNTCAGTTT 445
 QY 527 TGACTCCGCCATTACAGGTGATGCCAACCGCTGGATACAAAGTACAGAAATCACCA 586
 Db 446 TCCGGNACNTTTTACCGGATNTAATNGGGGTGGGAATGCNCA 496
 QY 587 AGCAGCAACATCTTACCGGAATCAGCATCGCTGGTGGAAAGTGGACCA 637

```

RESULT 7
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
cDNA clone 97SN1787, mRNA sequence.
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotas; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
REFERENCE 1 (bases 1 to 252)
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R, Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.I.
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
Unpublished (1998)

TITLE Oryza sativa
JOURNAL
COMMENT

Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhnamh@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers
1. 252
/organism="Oryza sativa"
/cultivar="Milyang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5'end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/cclone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN

Query Match 3.5%; Score 60; DB 12; Length 252;
Best Local Similarity 11.4%; Pred. No. 3,94e-61;
atches 25; Conservative 109; Mismatches 84; Indels 1; Gaps 1;

16 BAWMTTSYBCHGNBVWVCYVASHGNTMSYHNCTBRGTHCDCKNVKNSMTMTGTVNBNVSG 75
: : || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1341 GACCTTAATTCGGGCCCGCCCTAAGAAACTGGGTATGGGAAGCTTATAACACCCAT 1400
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 DWHYVBVNTKVDVGNHTRCSRWRBVRMAHYHDYTCBRYNNNDYHWHBBWYBTB7GCM 135
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1401 GGTGATTGCTGGTTGTCTGTGGTGGAGGACTCTCTTATTATTGATTATGTCAC 1460
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 TCTMWCWBHYNKCTASGWHSTNTNDYKSSNTWGTBYSYDKSMHGYWCBSBBVKYHTKVS 195
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1461 AATATTGAACCTGAATCAACCATGAAGGCACTACGTTACGTTAACTTGTCTTAACCTTGC 1520
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 TTRATRSYTCVRKYCYVMW-TRKKVVKYHVBBGCHBDT 233
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1521 TAGCTGGTTCGCTTCCTTGTGGGGGCAAGATGCAGTA 1559
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
cDNA clone 97SN1787, mRNA sequence.
ACCESSION AA754459
NID 92801165

```

KEYWORDS
SOURCE
ORGANISM

Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 252)
Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.Y., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
Unpublished (1998)

TITLE
JOURNAL
COMMENT

Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sunn20.asti.re.kr
Submitted by Baek He Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.

FEATURES
source
1..252
Location/Qualifiers
/organism="Oryza sativa"
/cultivar="Milyang23"
/note="vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone.lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN

Query Match 3.4%; Score 58; DB 12; Length 252;
Best Local Similarity 12.7%; Pred.No.1.52e-57;
Matches 29; Conservative 121; Mismatches 72; Indels 6; Gaps 6;

Db	10	RGCCCBAAWMTSYBCHGNBYVVCVASHGNYMSVNCTBRGTHCDC-KNVNWSTMT-WGT	67
Cp	525	GCCCATACATCGGGCTTGGATCCA-CAGCAAGCCAGCAGTCGCCGCAGACGATGTGA	467
Dd	68	VNVBNVSGDWHYWBNTKVDVGNHTRCSRWBRBVTRMAHYHDYTNGBNNNDYHMHBB	127
Cp	466	ACTCGGCGGACTCGGACGCCAGCACGCGGTAGACGATTGCCGCCATCATGACCGCAACG	407
Dd	128	MYBTGCMTCMCWBHYNTKTAGSWHTSTNYDKSNTTWGVTSYDKSMHGWCSSB	187
Cp	406	CCGACGAACGTC-CACATGATGT-GGTGGCCCTCTCTTCGAAGATCCCGGAC-CGCGAC	350
Dd	188	VKYHKVTSTRATRSYTCVRKYCVMMTKVKVKKYHVVBGBCHBDTSK	235
Cp	349	ATCTGTTCAAAGCCTCGGTAGTCCTTGAGATCTCCGACACTTCG	302

RESULT
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754458
NID 92801164
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 247)

AUTHORS Nahm, B.H., Kim, J.K., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P., Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y., Lee, M.C. and Eun, M.Y.

TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

JOURNAL Unpublished (1998)

COMMENT Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@n20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji University, Yongin, Korea. 449-728 bnhahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers
1. .247
/organism="Oryza sativa"
/cultivar="Milyang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2: XhoI; Directional cDNA library inserted into lambda ZAPII vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1784"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT 7 a 16 c 21 g 34 t 169 others
ORIGIN

Query Match 3.2%; Score 55; DB 12; Length 247;
Best Local Similarity 14.9%; Pred. No. 3.18e-52;
Matches 34; Conservative 106; Mismatches 85; Indels 3; Gaps 3;

Db 1 HWDCTMTVWRCGCCBAMN-KHFMHTBWCVRVCTTNNKNGHRTTWNDCSDN 59
Cp 374 CTTCGGAAGATCCCGGACGCGCATCTCTGTTAAAAAGCCCTCGGTAGCTTGAGAT 315
Db 60 AHCRTVBYARYSKYGYBTYSNVNDTNGTGTGKTTVNVHGNWNRCSNVVYVB 119
Cp 314 CTC-CGACACTTCGAAGT-CITGAGTAGTAGTACCGGTGAAGACGGATCCAGATGC 257
Db 120 TAYCYBHYBDRANVDDTRCTNDRGYNCTASDNGTSATKRVYDKTDSDCGGCGRK 179
Cp 256 CGCCACGCGCTCCCGGGTGTACGAATGAAGGGTCTCGGTGACGTCTCTGGCGCCAGA 197
180 VTGSSBVBRCGVNVVMTSMWTDKSTKMSMDMSRRSRVHYGRWB 227
Cp 196 CTGAGAGCGGACCTCGCGCGGGGTGCTCCGCGCCACCGAG 149

RESULT 10
LOCUS T42569 352 bp mRNA EST 07-JAN-1998
DEFINITION 5832 Lambda-PRL2 Arabidopsis thaliana cDNA clone 114CI17T, mRNA sequence.
ACCESSION T42569
NID 92757838
KEYWORDS EST
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryotes; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta;
eudicotyledons; Rosidae; Capparales; Brassicaceae; Arabidopsis.
1 (bases 1 to 352)
Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kendra, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E. and Somerville, C.
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729

REFERENCE 11
LOCUS T15209 597 bp mRNA EST 28-JUL-1995
DEFINITION crs698 Ricinus communis cDNA clone pcrs698 similar to cytochrome b5.
ACCESSION T15209
NID 9688663
KEYWORDS EST.
SOURCE castor bean clone-pcrs698 library-lambdaZAPST strain-Baker 296 vector-lambdaZAPII primer-T3 Rsite1-EcoRI Rsite2-XhoI Poly(A)+ RNA was purified from developing stage III to stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760, 1982) endosperm plus embryo of immature castor fruits. cDNA was synthesized and cloned into lambdaZAPII according to the instructions of the manufacturer (Stratagene); synthesis was primed from the poly(A) tail, and cloned directionally into XhoI (3') and EcoRI (5') sites. In few cases, sequence data indicated that this directionality was reversed. Partial cDNA clones predominate.
ORGANISM Ricinus communis
Eucaryotes; Embryophyta; Magnoliophyta; Magnoliopsida; Rosidae; Euphorbiales; Euphorbiaceae; Ricinus.
1 (bases 1 to 597)
van de Loo, F.J., Turner, S. and Somerville, C.
Expressed sequence tags from developing castor seeds
Plant Physiol. 108, 1141-1150 (1995)

COMMENT On Jan 8, 1998 this sequence version replaced gi:933288.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313t@clmsu.edu
Seq primer: T7.
Location/Qualifiers
1. .352
/organism="Arabidopsis thaliana"
/strain="var columbia"
/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."
/db_xref="taxon:3702"
/clone="114CI17T"
/clone_lib="Lambda-PRL2"
BASE COUNT 76 a 83 c 75 g 104 t 14 others
ORIGIN

Query Match 3.0%; Score 51; DB 11; Length 352;
Best Local Similarity 65.4%; Pred. No. 3.01e-45;
Matches 106; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Db 4 ATTCAAGGCAAGTCTACAAAGTCTCCNATTGGATTAAACTCATCCCGAGCGACACG 63
Qy 126 ATCCAGGCAAGTCTACAGCTCTCTCGTGGCGGAGCACCCTCGCGGAGGTC 185
Db 64 GTGATTCTAAATCTCGTTGGTCAAGACGTCAACGATGCTTTCATCGCATTTAATCCGGA 123
Qy 186 CGCTCTCTAGTCTGGCGGCGGAGGAGTCAACGACGCTTCATTCGTCGACCCGGGC 245
Db 124 ACCGTTGGCACCATCTCGACCATCTTTTCACCGGTTACAC 165
Qy 246 ACGGCGTGGCGATCTGATCGCTCTTCACGCGGTACTAC 287

Contact: Somerville CR
Carnegie Institution
Carnegie Institution, 290 Panama St., Stanford, CA 94305
Tel: 4153251521
Email: crs@andrew.stanford.edu.

FEATURES

source

1. .597
Location/Qualifiers
/organism="Ricinus communis"
/clone="pers98"
/strain="Baker 296"

BASE COUNT 167 a 111 c 120 g 171 t 28 others

ORIGIN

Query Match 2.9%; Score 50; DB 16; Length 597;
Best Local Similarity 62.8%; Pred. No. 1.59e-43;
Matches 130; Conservative 0; Mismatches 76; Indels 1; Gaps 1;

Db 148 GAAACAAAGATACATACAAAAGAGATCTTGAAGACCAATAACCCCTGGAGATCTC 207

QY 57 GAAGCTAGAGATATATACACGGCGGAGGACCTCCGCCCCACACACAGTCCGGGATCTC 116

208 TGGNTCTCAATACAGGCAAGATCTACAAATGTTACTGATGGGCAAGACCATCTGCT 267

QY 117 TGGATCTCCATCCAGGCAAGGTCTACAGCTGCTCTCGTGGCGGCGGAGCACCCTCGC 176

Db 258 GGTGTGTCGCCCTTGCYTCACTTGTCTGGCCCAAGATGTTACTGATGCTTTTGTGCTTA 327

QY 177 GCGAGAGTCCCGGCTCCTCAGTCTGCGCGCGGAGGACGCTCAGCGCTTCATT-GCGTA 235

328 CCATCTCGGAGTCTGTTGCAATATCT 354

QY 236 CCACCGGCGACGGCTGGCGGCTCT 262

RESULT 12

LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
cDNA clone 97SN1784, mRNA sequence.

ACCESSION AA754458

NID 92801164

KEYWORDS EST.

SOURCE rice.

ORGANISM

Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.

1 (bases 1 to 247)

Nahm,B.H., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,

Kim,W.T., Kim,W.Y., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

Lee,M.C. and Eun,M.Y.

Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

Unpublished (1998)

JOURNAL

COMMENT

Contact: Eun M.Y.

Department of Cytogenetics

National Inst. of Agri. Sci. and Tech, RDA

Suwon, Kyunggido, Korea

Tel: 82 331 290 0301

Fax: 82 331 290 0307

Email: myeun@sun20.osti.re.kr

Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji

University, Yongin, Korea. 449-728 bhnam@bioserver.myongji.ac.kr

Seq primer: M13 Reverse Primer.

FEATURES

source

1. .247
Location/Qualifiers

/organism="Oryza sativa"

/cultivar="Milyang23"

/note="Vector: pBluescript SK(+); Site1: EcoRI; Site2:

XhoI; Directional cDNA library inserted into lambda ZAPII

vector at 5' end with EcoRI and 3' end with Xho I site."

/db_xref="taxon:4530"

/clones="97SN1784"

/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"

/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others

ORIGIN

Query Match 2.8%; Score 48; DB 12; Length 247;
Best Local Similarity 15.1%; Pred. No. 4.09e-40;
Matches 35; Conservative 103; Mismatches 91; Indels 3; Gaps 3;

Db 3 DCTMNTVWRCGCCBAWNNKHTHMTBWCVRVRYGTTTNGKHNGRTTTWNDCSDNAHC 62

QY 667 TCGACTAGACCCCGACCTCCAGCACATCCCGTATTCGCGCTCCACCCGACTCTCA 726

Db 63 RYTVBYIYARSKYGYGTBYISNNVDNTGTVGKTVVNVHSGWNNRCSNVVYVWBRAY 122

QY 727 ACTCCATCACCTCG-GTCTTCTATGCGCGAGTCTCTGAAATTCGACGAAAGTGGCAGGTTC 785

Db 123 CDYBHYBDRANHVDDTRCTNDRGYCNYTASDNGTSATRVYDGYDKTDSDCGGGCKRWKTYI 182

QY 786 CTAGTCAGCTACCAGCACTGGACCTACTACCGGCTCATGATCTTCGCGCGAGTCAACCTC 845

Db 183 GSSBYBRGCVNVMVTRTSMWTDKSTKMSMDMSRRSRVHYGRWMBNKKRGMS 234

QY 846 TTCAIC-C-AGACCTTTTATGCTCTCACACGCGGCGAGCTCCCTGACCG 895

RESULT 13

LOCUS AF034173 2275 bp mRNA EST 22-DEC-1997
DEFINITION Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.
ACCESSION AF034173

NID 92707735

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 2275)

Tripodis,N. and Ragoussis,J.

Generation of a transcription map in the region immediately

centromeric to human MHC across the 6p21.2-6p21.3 chromosomal

boundary

Unpublished

2 (bases 1 to 2275)

Tripodis,N. and Ragoussis,J.

Direct Submission

Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,

Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK

Location/Qualifiers

1. .2275

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="6p21.3"

/clones="ntcon2 contig"

/tissue_type="liver; brain"

/dev_stage="fetus"

/note="similar to Brl40"

BASE COUNT 438 a 619 c 470 g 599 t 149 others

ORIGIN

Query Match 2.7%; Score 46; DB 11; Length 2275;

Best Local Similarity 15.6%; Pred. No. 9.56e-37;

Matches 25; Conservative 83; Mismatches 50; Indels 2; Gaps 2;

Db 1492 KRMTGMYKRYRAMMAMCAMMACWYWKRMGMKKWKRYKRYKRYKRYKRYKRYKRY 1551

QY 1444 TATTGTTTATGTCACAAATATTGAAC-TAATAACCATGAAGGACACTACCTTACGCTT 1502

Db 1552 TTYTGWCCCTSKASACAMRWGYSRSSRYSWYWGWSGSCYGMKRYRYRYRY 1611

QY 1503 AACTTTGCTTAACCTTGTAGCTGTTGCTTCCCTTGTGGGGCAAGTGCAGTATTT 1562

Db 1612 WTKWT-TWYMWSPRWTMTTYYTWTWTWTWTWTWTWTWT 1650
 QY 1563 ATTTCCTATCCCATGCTATTTTGTGATTATGTTCTATT 1602

RESULT 14
 LOCUS A1107749 448 bp mRNA EST 26-AUG-1998
 DEFINITION GH05586.5prime GH Drosophila melanogaster head pot2 Drosophila
 melanogaster cDNA clone GH05586 5prime, mRNA sequence.

ACCESSION A1107749
 NID 93475402

KEYWORDS EST.
 SOURCE fruit fly.

ORGANISM Drosophila melanogaster

REFERENCE 1 (bases 1 to 448)
 AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
 Brokstein,P., Lewis,S. and Rubin,G.M.

BDGP/HMI Drosophila EST Project

UNpublished (1997)

COMMENT

Contact: Harvey, D.

G. M. Rubin-Molecular and Cell Biology

University of California Berkeley

539 LSA, Berkeley, CA 94720-3200, USA

Fax: 510 643 9947

Email: http://fruitfly.berkeley.edu/EST, est@fruitfly.berkeley.edu

Plate: 55 row: H column: 2

High quality sequence stop: 366.

Location/Qualifiers

1..448

/organism="Drosophila melanogaster"

/note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:

pot2; Sized fractionated cDNAs were directly ligated into

pot2. Plasmid cDNA library."

/db_xref="taxon:7227"

/clone="GH05586"

/clone_lib="GH Drosophila melanogaster head pot2"

/sex="male and female"

/dev_stage="adult"

/lab_host="DH5 - alpha"

BASE COUNT 111 a 120 c 126 g 91 t

ORIGIN

Query Match 2.0%; Score 34; DB 17; Length 448;

Best Local Similarity 62.5%; Pred. No. 1.05e-17;

Matches 85; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 52 CACAACAAGCCACGGATCTGTGGTGTGTCATCGACAACAAGGTCTACGATGTGACCAAG 111

QY 96 CACAACAAGTCGGCGATCTCTGGATCTCCATCCAGGCAAGGTCTACGATGTGCTCGG 155

Db 112 TTCGCTCTCGAGCATCCCGTGGCGGAGGAATCCCTGGTGGATGAGCGCGGTCCGATGCC 171

QY 156 TGGCGCGGAGACACCCCGCGGAGGTCCCGTCTCAGTCTGCGCGGAGGACGTC 215

Db 172 ACCAAGGCTTCAATG 187

QY 216 ACCGACGCTTCATTG 231

Search completed: Thu Feb 18 17:40:14 1999

Job time : 4650 secs.

RESULT 15
 LOCUS A1107703 613 bp mRNA EST 26-AUG-1998
 DEFINITION GH05526.5prime GH Drosophila melanogaster head pot2 Drosophila
 melanogaster cDNA clone GH05526 5prime, mRNA sequence.

ACCESSION A1107703
 NID 93475356

KEYWORDS EST.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

REFERENCE 1 (bases 1 to 613)
 AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
 Brokstein,P., Lewis,S. and Rubin,G.M.

BDGP/HMI Drosophila EST Project

UNpublished (1997)

COMMENT

Contact: Harvey, D.

G. M. Rubin-Molecular and Cell Biology

University of California Berkeley

539 LSA, Berkeley, CA 94720-3200, USA

Fax: 510 643 9947

Email: http://fruitfly.berkeley.edu/EST, est@fruitfly.berkeley.edu

Plate: 55 row: C column: 2

High quality sequence stop: 609.

Location/Qualifiers

1..613

/organism="Drosophila melanogaster"

/note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:

pot2; Sized fractionated cDNAs were directly ligated into

pot2. Plasmid cDNA library."

/db_xref="taxon:7227"

/clone="GH05526"

/clone_lib="GH Drosophila melanogaster head pot2"

/sex="male and female"

/dev_stage="adult"

/lab_host="DH5 - alpha"

BASE COUNT 173 a 154 c 150 g 136 t

ORIGIN

Query Match 2.0%; Score 34; DB 17; Length 613;

Best Local Similarity 62.5%; Pred. No. 1.05e-17;

Matches 85; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 52 CACAACAAGCCACGGATCTGTGGTGTGTCATCGACAACAAGGTCTACGATGTGACCAAG 111

QY 96 CACAACAAGTCGGCGATCTCTGGATCTCCATCCAGGCAAGGTCTACGATGTGCTCGG 155

Db 112 TTCGCTCTCGAGCATCCCGTGGCGGAGGAATCCCTGGTGGATGAGCGCGGTCCGATGCC 171

QY 156 TGGCGCGGAGACACCCCGCGGAGGTCCCGTCTCAGTCTGCGCGGAGGACGTC 215

Db 172 ACCAAGGCTTCAATG 187

QY 216 ACCGACGCTTCAATG 231

Search completed: Thu Feb 18 17:40:14 1999

Job time : 4650 secs.

MAEPLH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsearch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Date: Thu Feb 18 17:50:01 1999; MasPar time 122.19 Seconds
Tabular output not generated. 1204.694 Million cell updates/sec

Title: >US-08-934-254-26
Description: (1-1702) from US08934254.seq
Perfect Score: 1702
N.A. Sequence: 1 CCCCAAAATTTTCATGTT.....TTTTGGTAAAAA... 1702
Comp: GGGGTTTTTAAAGTAACAA.....AAAAACATTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 165359 seqs, 43243793 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:5A_COMB 2:5B_COMB 3:5C_COMB 4: PCT9_COMB 5:backfiles1

Statistics: Mean 9.108; Variance 5.151; scale 1.768

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	172	10.1	1685	1	US-08-366- Sequence 4, Applicatio	5.25e-102
2	172	10.1	1685	3	US-08-789- Sequence 4, Applicatio	5.25e-102
3	91	5.3	7218	2	US-08-232- Sequence 14, Applicati	5.86e-44
4	46	2.7	9653	3	US-08-388- Sequence 22, Applicati	7.85e-14
5	40	2.4	9653	3	US-08-388- Sequence 22, Applicati	3.99e-10
6	37	2.2	7218	2	US-08-232- Sequence 14, Applicati	2.52e-08
7	36	2.1	2151	1	US-08-238- Sequence 5, Applicatio	9.86e-08
8	34	2.0	2151	1	US-08-238- Sequence 5, Applicatio	1.45e-06
9	27	1.6	74	4	PCT-US95-1 Sequence 94, Applicati	1.16e-02
10	28	1.6	81	4	PCT-US95-1 Sequence 92, Applicati	3.36e-03
11	27	1.6	242	2	US-08-273- Sequence 1, Applicatio	1.16e-02
12	26	1.5	74	4	PCT-US95-1 Sequence 94, Applicati	3.89e-02
13	25	1.5	74	4	PCT-US95-1 Sequence 100, Applicati	1.28e-01
14	26	1.5	75	4	PCT-US95-1 Sequence 99, Applicati	3.89e-02
15	26	1.5	81	4	PCT-US95-1 Sequence 92, Applicati	3.89e-02
16	25	1.5	81	4	PCT-US95-1 Sequence 98, Applicati	1.28e-01
17	25	1.5	82	4	PCT-US95-1 Sequence 97, Applicati	1.28e-01
18	26	1.5	82	4	PCT-US95-1 Sequence 97, Applicati	3.89e-02
19	25	1.5	92	3	US-08-353- Sequence 16, Applicati	1.28e-01
20	25	1.5	242	2	US-08-273- Sequence 1, Applicatio	1.28e-01

21	25	1.5	2016	1	US-08-294- Sequence 1, Applicatio	1.28e-01
22	25	1.5	2016	4	PCT-US95-0 Sequence 2, Applicatio	1.28e-01
c 23	26	1.5	5852	1	US-07-867- Sequence 144, Applicat	3.89e-02
c 24	23	1.4	66	1	US-08-471- Sequence 142, Applicat	1.31e-00
25	23	1.4	69	1	US-08-471- Sequence 142, Applicat	1.31e-00
c 26	23	1.4	69	1	US-08-471- Sequence 142, Applicat	1.31e-00
27	23	1.4	74	4	PCT-US95-1 Sequence 100, Applicati	1.31e-00
28	24	1.4	75	4	PCT-US95-1 Sequence 99, Applicati	4.14e-01
29	23	1.4	76	3	US-08-478- Sequence 95, Applicati	1.31e-00
30	23	1.4	76	1	US-08-120- Sequence 95, Applicati	1.31e-00
31	23	1.4	81	4	PCT-US95-1 Sequence 30, Applicati	1.31e-00
32	23	1.4	90	3	US-08-442- Sequence 6, Applicatio	4.14e-01
33	24	1.4	165	4	PCT-US93-0 Sequence 1, Applicatio	4.14e-01
34	24	1.4	1662	3	US-08-651- Sequence 1, Applicatio	4.14e-01
35	23	1.4	2830	4	PCT-US93-0 Sequence 1, Applicatio	1.31e-00
36	23	1.4	2830	1	US-07-882- Sequence 10, Applicatio	1.31e-00
c 37	23	1.4	5181	3	US-08-257- Sequence 25, Applicati	4.01e-00
c 38	22	1.3	84	4	PCT-US95-1 Sequence 120, Applicat	4.01e-00
c 39	22	1.3	84	2	US-08-133- Sequence 4, Applicatio	4.01e-00
c 40	22	1.3	91	4	PCT-US95-0 Sequence 1, Applicatio	4.01e-00
c 41	22	1.3	768	3	US-08-567- Sequence 10, Applicati	4.01e-00
42	22	1.3	2088	1	US-08-073- Sequence 3, Applicatio	4.01e-00
43	22	1.3	2504	2	US-08-483- Sequence 1, Applicatio	4.01e-00
44	22	1.3	2626	1	US-08-156- Sequence 1, Applicatio	4.01e-00
45	22	1.3	49377	2	US-08-764- Sequence 1, Applicatio	4.01e-00

ALIGNMENTS

RESULT 1
ID US-08-366-779-4 STANDARD; DNA; UNC; 1685 BP.
AC xxxxxx

Sequence 4, Application US/08366779
Sequence 4, Application US/08366779
Patent No. 5614393

GENERAL INFORMATION:

APPLICANT: Thomas, Terry L.
APPLICANT: Reddy, Avutu S.
APPLICANT: Nuccio, Michael
APPLICANT: Freytsinet, Georges L.
APPLICANT: Nunberg, Andrew N.
TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/366,779
FILING DATE: 30-DEC-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Presser, Leopold
REGISTRATION NUMBER: 19,827
REFERENCE/DOCKET NUMBER: 8383ZYXW
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
SEQUENCE CHARACTERISTICS:
LENGTH: 1685 base pairs
TYPE: nucleic acid

CC STRANDEDNESS: both
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (c
SQ SEQUENCE 1685 BP; 431 A;

Query Match 10.1%; Score 172; DB 1; Length 1685;
Best Local Similarity 57.7%; Pred. No. 5.25e-102;
Matches 697; Conservative 0; Mismatches 507; Indels

Local similarity 97.7%, freq. no. 3,256 102;
 Matches 697; Conservative 0; Mismatches 507;
 Indels 3; Gaps 2;

Db	42	CAATGGCTGCTCAATCAAGAAATACATTACCTCAGATGAATCAAGAACACGAGATAAAC	101
QY	43		102
	46	CAATGAGGGCGAAGCTTAAGAAGTATATCACGGCGGAGGACCTCCGCGCCACAACAAGT	105
Db	102	CGGAGATCTATGGATCTCGATTCAAGSGAAAGCCATGATGTTTCGGATTGGGTGAAGA	161
	106		162
QY	106	CGGGCGATCTCTGGATCTCCATCCAGGGCAAGGCTCTACGACTGCTCTCGGTGGCGGGGG	165
	162	ACCATCCAGGTGGCAGCTTTCCTTGAACAGTCTTCTGCTGCTCAAGAGGTAACTGATGCAT	221
QY	166	AGCACCCCGCGCGAGGTCCCGTCTCAGTCTTGCCGCGCCAGGACGTCACCGAGCGCT	225
	222	TTGTTGCATTCCATCTCGCTCTACATGGAAGAACTCTGATAAGTATTTTCACTGGGTATT	281
Db	226	TCATTCGGTACCACCGCGCACGGCTGGCGCATCTGGATCCGCTCTTCACCGGCTACT	285
	282	A-T--CTTAAAGATACTCTGTTCTCGAGTTTCTAAAGATTATAGGAAGCTTGTGTTG	338
QY	286	ACTATCCTCAAGACTTCGAAGTGTGGAGATCTCCAAGGACTACCGGAGGCTTTTGAACG	345
	339	AGTTTCTAAATGGTTTTGTATGACAAAAAGGTCAATATTAAGTTTGCACACTTTGTGCT	398
QY	346	AGATGTCGCGGTCCGGGATCTTCGAGAAGAGGGCCACACATCATGTGGACGTTTCGTCG	405
	399	TTATAGCAATGCTGTTTGTATAGTGTTTATAGGGTTTTTGTGTTGTGAGGGTGTTTGG	458
QY	406	GGTTCGGGTCTATGATGGCGCAATCGTCTACGGCGTCTGCGCTCGGAGTCGTCGGAG	465
	459	TACATTGTTTTCTGGGTGTTTGATGGGGTGTCTTTTGGATTCCAGAGTGGTTGGATTGGAC	518
QY	466	TTACATGCTCTCGCGCGCACTGCTGGGCTTCTGTGATCCAACCGCGTATGTGGGCC	525
	519	ATGATGCTGGGCATTATATGTTAGTGTCTGATTCAAGGCTTAATAAGTTTATGGGTATTT	578
QY	526	ATGATCCGGCCATTACCAAGGTGATGCCAACCCGTGGATACAACAGATCACGCAACTCA	585
	579	TTGCTGCAAAATGCTTTCAGGAATAAGTATTGGTTGGTGGAAATGGAACCAATATGCAC	638
QY	586	TAGCAGGCACAATCCTAACCGGAATCAGCATCGCGTGGTGGAAAGTGAGCCACACGCC	645
	639	ATCACATTCCCTGTAAATAGCCTTGAATATGACCCCTGATTTCACAATATATACCATTCCTTG	698
QY	646	ACCACCTGCCCTGCAACAGCCTCGACTACGACCCGCGACTCCAGCACATCCCGCTATTTCG	705
	699	TTGTGTCTCCAAAGTTTTTGGTTTCACCTCACCTCTCATTTCTATCAGAAAAGTTTGACTT	758
QY	706	CGTCTCCACCGCACTTTCAACTCCATCACCTCGGCTCTTATGGCGAGTCTCTGAAT	765
	759	TTGACTCTTTATCAAGATCTTTGTAAGTATCAACATTGGACATTTTACCCCTATTATGT	818
QY	766	TCGACGAAGTGGCAGGTTCTTAGTCAGTACCAGCACTGGACCTACTACCCGGTCAATGA	825
	819	GTGCTGTAGGCTCAATATGTAATGATACAACTCTCATAAATGTTGTGGACCAAGAAATG	878
QY	826	TCTTGGCGCGAGTCAACCTCTTCATCCAGACCTTTTATTGCTCTCATCCAGGCGCAGC	885
	879	TGTCCTATCGACTCAGGAACHTCTGGGATGCCTAGTGTCTCTGATTTGGTACCGGTTCG	938
QY	886	TCCTTGACCGCGCTCTAACTTAATGGGTATCGCGGTTTTCTGGACGTGTGTTCCGCTCT	945
	939	TGTTTCTGTTTGCTTAATGGGGTGAAGAAATATATGTTTGTATTTCAGATTATACAG	998
QY	946	TCGTATCTGTCTCCGAACCTGGCTGAACGGTTCGGGTCTGTCCTCATCAGCTTTGGG	1005

Db	999	TGACTGGAATGCAACAAGTTCAGTCTCTCTTGAACCACTTCTCTTCAAGTGTATATGTTG	1058
Qy	1006	TCACGGCGATCCACGACGTCCTCAGTTTCACGCTCAACCACTTCTCCGGCGACACATACGTGG	1065
Db	1059	GAAGCCTTAAGGGAATAATTGGTTTGAGAAACAACGATGGGACACTTCACATTTCTT	1118
Qy	1066	GCCCCCAGGCGCACTGGTTCGAGAACGACAGAAAGGAGCAGATCGATATCAGCT	1125
Db	1119	GTCTCTCTTGGATGGATGGTTTTCATGGTGGATTCGAAATCCAAATTGAGCATCATTTGT	1178
Qy	1126	GCCACCGTGGATGGACTGGTCTTTGGTGGCTCGAGTTCGAGTTGGAGCACCACCTTGT	1185
Db	1179	TTCCCAGATGCGCTAGATCAACCTTAGAANAATCTCGCCCTAGTGATCAGTATTGCA	1238
Qy	1186	TCCCTAGCTGCCCGCTGGCAGCTTAGAAGATGGCCCTTGGCTCGGACATTGTGTA	1245
Db	1239	AGAAACA	1245
Qy	1246	AGAAGA	1252

RESULT	2
ID	US-08-1
AC	XXXXXX

RECORD #	ID	US-08-789-936-4	STANDARD	DNA	UNC	1685 BP

```

AC      xxxxxx
DE
DT
CC      Sequence 4, Application US/08789936
CC      Sequence 4, Application US/08789936
CC      Patent No. 5789220
CC      GENERAL INFORMATION:
CC      APPLICANT: Thomas, Terry L.
CC      APPLICANT: Reddy, Avutu S.
CC      APPLICANT: Nuccio, Michael
CC      APPLICANT: Freyssinet, Georges L.
CC      APPLICANT: Nunberg, Andrew N.
CC      TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A
CC      TITLE OF INVENTION: DELTA 6-DESATURASE
CC      NUMBER OF SEQUENCES: 25
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Scully, Scott, Murphy & Presser
CC      STREET: 400 Garden City Plaza
CC      CITY: Garden City
CC      STATE: New York
CC      COUNTRY: United States
CC      ZIP: 11530
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patentin Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/789,936
CC      FILING DATE: 28-JAN-1997

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/366,779
FILING DATE: 30-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Presser, Leopold
REGISTRATION NUMBER: 19,827
REFERENCE/DOCKET NUMBER: 83832YXW
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1685 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE 1685 BP: 431 A; 277 C; 357 G; 620 T; 0 OTHER.

Query Match 10.1%; Score 172; DB 3; Length 1685;
Best Local Similarity 57.7%; Pred. No. 5.25e-102;
Matches 697; Conservative 0; Mismatches 507; Indels

Db	42	CAATGCTGCTCAAAATCAAGAATACATTTACCTCAGATGAATCAAGAACACACGATAAAC	101
QY	43		102
	46	CAATGGAGGGCGAAAGCTAAAGAAGTATATCACGGCGGAGGACCTCCGCCGCCACAACAAGT	105
Db	102	CGGGACATATGGATCTCGATTCGAAGGGAAGACCTATGATGTTTCGGATGCGGTGGGAAG	161
QY	106		162
	162	ACCATCCAGGTGGCAGCTTTCCCTTGAAGAGTCTTGCTGGTCAAGAGGTAACATGATGCAT	221
QY	166	AGCACCCCGGGCGGAGTGCCCGCTCCTCAGTCTGCGCGCCAGGACGTCACCGACGCT	225
Db	222	TTGTTGCATTCCATCCTGCCTCTACATGGAAGAATCTTGATAAGTTTTTCACTGGGTATT	281
QY	226	TCATTGGCGTACCACCCGGCGACGGGTGCGGCATCTGGATCCGCTCTTCACCGGCTACT	285
	282	A-T--CTTAAAGATTAATCTGTTCTTCGAGAGTTTCTAAAGATTATAGGAAGCTTGTGTTTG	338
QY	286	ACTACCTCAAGACTTCGAAGTGTCCGAGATCTCCAAAGGACTACCGAGGCGTTTGAAGC	345
Db	339	AGTTTTCTAAATGGGTTTGATGACAAAAGGTCAATATATGTTTGCAACTTTGTGCT	398
QY	346	AGATGTCGGGTCCGGGATCTTCGAGAAAGAGGGCCACCACATCAATGTGGAGCTTCGTGC	405
	399	TTATAGCAATGCTGTTTGCTATGAGTGTTATATGGGGTTTTGTTTTGTAGGGTGTTTTGG	458
QY	406	GC GTTGGGTGATGATGGCGCAATCGTCTACGGCGTCTGCGGTGCGAGTCCGTCGGAG	465
Db	459	TACATTTGTTTTTGGGTGTTTGATGGGGTTTCTTTTGGATTCAGAGTGGTTGGATTGGAC	518
QY	466	TTCAATGCTCTCGCGCGCACGCTGGGCTTGCTGTGGATCCAAGCCGCTATGTGGGCG	525
Db	519	ATGATGCTGGGCATTAATATGTAGTCTCTGATTCAGGCTTAAATAGTTTATGGGTATT	578
QY	526	ATGACTCCGGCCATACCAAGGTGATCCCAACCCGTGGATACACAGAAATCAGCAACTCA	585
Db	579	TTGCTGCAAAATGCTCTTCAGGAATAAGTATTGGTTGGTGGAAATGGAAACCAATATAGT	638
QY	586	TACAGCGAACATCCTAACCGGAATCAGATCGGTGGTGGAGTGGACCCCAACGCGCC	645
Db	639	ATCACATTGCTCTAATAGCTTGAATATGACCCCTGATTTACAATATATACCATCTCTTG	698
QY	646	ACCACCTGCCCTGCACAGCGCTCGACTACGACCCCGACCTCCAGCACATCCCGGTATT	705
	699	TTGTGTTCCCAAGTTTTTTGGTTCACCTCACCTCTCATTTCTATGAGAAAAGTTGACT	758
QY	706	CCGTCTCCACCCGACTCTTCACTCCATCACTCCGCTCTCTATGSCCGAGTCCCTGAAT	765
Db	759	TTGACTCTTTATCAAGATTCTTTGTAAGTATCAACATTGGACATTTTACCCTATTATGT	818
QY	766	TCGACGAAGTGGCAGCGTTCCCTAGTCAGCTACACGACTGGACCTACTACCCGGTCATGA	825
Db	819	GTGCTGCTAGGCTCAATATGTATGTACAAATCTCTCATATGTTTGTGACCAAGAGAAATG	878
QY	826	TC TTGCGCGAGTCAACCTTCTCATCCAGACCTTTTATGTCTCTCACAGCGCGGACG	885
Db	879	TGTCCTATCGAGCTCAGGAACCTTTGGGATGCCGTAGTGTCTCGATTTTGGTACCCGTTGC	938
QY	886	TCCTTGACCGCGCTCTAAACTTAATGGGTATCGCGGTTTTCTGGACGTGGTTCCCGCTCT	945
Db	939	TTGTTTCTTTGTTTGGCTTAATGGGGTGAAGAATTAATGTTTATATGCAAGTTTATCAG	998
QY	946	TCGTATCTTTGCTCCCGGAATGGCCTGGAACGGTTCGGGTTTCGTCCTCATCAGCTTTCGG	1001
Db	999	TGACTGGATGCAACAAGTTCAGTTCTCTTCTTGAACCACCTCTCTTCTCAAGTGTTTATGTTG	1051
QY	1006	TCAGGGGATCCAGCAGCTCCAGTTCACGCTCAACCACCTTCTCCGCGGACACATAGCTGG	1061
Db	1059	GAAAGCCTTAAAGGAATAATTTGGTTTGAGAAAACAAACGGATGGACACTTGCATTTCTTT	1111

[illegible]

RESULT	3	
ID	US-08-232-463-14	STANDARD; DNA; UNC; 7218 BP.
AC	xxxxxx	
DT		
DE	Sequence 14, Application US/08232463	
CC	Sequence 14, Application US/08232463	
CC	Patent No. 5670367	
CC	GENERAL INFORMATION:	
CC	APPLICANT: DORNER, F.	
CC	APPLICANT: SCHEIFLINGER, F.	
CC	APPLICANT: FALKNER, F. G.	
CC	TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS	
CC	NUMBER OF SEQUENCES: 52	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: Foley & Lardner	
CC	STREET: 1800 Diagonal Road, Suite 500	
CC	CITY: Alexandria	
CC	STATE: VA	
CC	COUNTRY: USA	
CC	ZIP: 22313-0299	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: PatentIn Release #1.0, Version #1.251	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: US/08/232,463	
CC	FILING DATE:	
CC	CLASSIFICATION: 435	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER: US/07/935,313	
CC	FILING DATE:	
CC	APPLICATION NUMBER: EP 91 114 300.6	
CC	FILING DATE: 26-AUG-1991	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME: BENT, Stephen A.	
CC	REGISTRATION NUMBER: 29,768	
CC	REFERENCE/DOCKET NUMBER: 30472/114 IMMU	
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE: (703)836-9300	
CC	TELEFAX: (703)683-4109	
CC	TELEX: 899149	
CC	INFORMATION FOR SEQ ID NO: 14:	
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH: 7218 base pairs	
CC	TYPE: nucleic acid	
CC	STRANDEDNESS: single	
CC	TOPOLOGY: linear	
CC	IMMEDIATE SOURCE:	
CC	CLONE: pTZgpt-F15	
CC	SEQUENCE 7218 BP: 1944 A; 1491 C; 1486 G; 1929 T; 365	

```
Query Match      5.3%; Score 91; DB 2; Length 7218;  
Best Local Similarity 0.3%; Pred. No. 5.86e-44;  
Matches       1; Conservative    229; Mismatches 139; Indels   0; Gaps   0;
```

```

QY 634 CCCAACGCCACCACCTCGCTGCAACAGCTCGACTACGACCGGACCTCCAGACA 693
Db 1127 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1186
QY 694 TCCCGGATTCGGCGTCCACCGGACTCTCAACTCCATCCTCGTCTTATGCGC 753
Db 1187 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1246
QY 754 GAGTCTGAATTCGAGGAGTGCAGGCTCTAGTACGCTACGACCTGACCTACT 813
Db 1247 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1306
QY 814 ACCGGTCATGATCTCGCGGAGTCAACCTCTTCATCCAGACCTTTTATGCTCTCA 873
Db 1307 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1366
QY 874 CCAGCGCGGAGCTCCCTGACCGGCTCTAAACTAATGATGCTCGGCTTCTGACGT 933
Db 1367 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1426
QY 934 GTTCGGCTCTCGTATCTGTCTCCGAACTGGCTGAACGTTGCGGTTCTGCTCTCA 993
1427 YYYYYYYG 1435
QY 994 TCAGCTTTG 1002
```

```

RESULT 4
ID US-08-388-672A-22 STANDARD; DNA; UNC; 965 BP.
AC xxxxxx
DT
DE Sequence 22, Application US/08388672A
CC Sequence 22, Application US/08388672A
CC Patent No. 5795961
CC GENERAL INFORMATION:
CC APPLICANT: Wallace, T. Paul
CC APPLICANT: Harris, William J.
CC APPLICANT: Carr, Frank J.
CC APPLICANT: Old, Lloyd J.
CC APPLICANT: Well, Sydney
CC APPLICANT: Kitamura, Kunio
CC TITLE OF INVENTION: Recombinant Human Anti-Lewis B
CC TITLE OF INVENTION: Antibodies
CC NUMBER OF SEQUENCES: 25
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Felfe and Lynch
CC STREET: 805 Third Avenue
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/388,672A
CC FILING DATE: 14-FEB-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hanson, No. 5795961man D.
CC REGISTRATION NUMBER: 30,946
CC REFERENCE/DOCKET NUMBER: LUD 5409
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-688-9200
CC TELEFAX: 212-838-3884
CC INFORMATION FOR SEQ ID NO: 22:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 965 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: unknown
CC MOLECULE TYPE: DNA (genomic)
QY SEQUENCE 965 BP; 192 A; 170 C; 226 G; 200 T; 177 OTHER.
```

```

CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 965 BP; 192 A; 170 C; 226 G; 200 T; 177 OTHER.

Query Match 2.7%; Score 46; DB 3; Length 965;
Best Local Similarity 17.4%; Pred. No. 7.85e-14;
Matches 26; Conservative 76; Mismatches 44; Indels 3; Gaps 3;

Db 786 VHYSGVSRSTSTASDYTTSYWGVGRGWGDYGGGYTNYNKGRTVTMADTSSNRSS 845
QY 104 GTCGCGGATCTCTGATCTCCATCCAGGCAAGTCTACGACTGCTCTCGTGGGGCGC 163
Db 846 VTAAATVYVCVRGRSYDSDGDYWGTTVTYVSSHUVDKMTSSSSASVGDRTVTCRSST 905
QY 164 GGA-GCACC CGCGGGAGGTCCCGCTCTCACTGCGC-CGGC-CAGGACGTCACCGA 220
Db 906 HGNGTYYWYKGRKRYVSNRSGVSRSGS 934
QY 221 CGCTTCATTGCGTACCAACCGCGGCACGG 249

RESULT 5
ID US-08-388-672A-22 STANDARD; DNA; UNC; 965 BP.
AC xxxxxx
DT
DE Sequence 22, Application US/08388672A
CC Sequence 22, Application US/08388672A
CC Patent No. 5795961
CC GENERAL INFORMATION:
CC APPLICANT: Wallace, T. Paul
CC APPLICANT: Harris, William J.
CC APPLICANT: Carr, Frank J.
CC APPLICANT: Old, Lloyd J.
CC APPLICANT: Well, Sydney
CC APPLICANT: Kitamura, Kunio
CC TITLE OF INVENTION: Recombinant Human Anti-Lewis B
CC TITLE OF INVENTION: Antibodies
CC NUMBER OF SEQUENCES: 25
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Felfe and Lynch
CC STREET: 805 Third Avenue
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/388,672A
CC FILING DATE: 14-FEB-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hanson, No. 5795961man D.
CC REGISTRATION NUMBER: 30,946
CC REFERENCE/DOCKET NUMBER: LUD 5409
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-688-9200
CC TELEFAX: 212-838-3884
CC INFORMATION FOR SEQ ID NO: 22:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 965 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 965 BP; 192 A; 170 C; 226 G; 200 T; 177 OTHER.

Query Match 2.4%; Score 40; DB 3; Length 965;
Best Local Similarity 17.1%; Pred. No. 3.99e-10;
Matches 30; Conservative 77; Mismatches 67; Indels 1; Gaps 1;
```

[illegible]

RESULT 6
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx

Sequence 14, Application US/082322463
Sequence 14, Application US/082322463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFELINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
City: Alexandria

COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149

CC INFORMATION FOR SEQ ID NO: 14:
CC
CC SEQUENCE CHARACTERISTICS:
CC
CC LENGTH: 7218 base pairs
CC
CC TYPE: nucleic acid
CC
CC STRANDEDNESS: single
CC
CC TOPOLOGY: linear
CC
CC IMMEDIATE SOURCE:
CC
CC CLONE: pTZqpt-F15

SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

```
Query Match      2.2%; Score 37; DB 2; Length 7218;
Best Local Similarity 0.8%; Pred. No. 2.52e-08;
Matches 3; Conservative 197; Mismatches 163; Indels 0; Gaps 0;
```

Db 1064 GATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1123
|||||::: : : : : : : : : :
Cp 1436 GATCCTCAACCCACACACAAAACCACATCGGTGTTATAAGCTTCCCATA 1377

[illegible]

RESULT 7
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx

Sequence 5, Application US/08238163
Sequence 5, Application US/08238163
Patent No. 5569830
GENERAL INFORMATION:

CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: SPOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS
CC TITLE OF INVENTION: POLYGALACTIC ACID
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC
CC STREET: Stewart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC
CC COUNTRY: US
CC ZIP: 94105-1493

CC ZLF: 9410J-1493
CC COMPUTER READABLE FORM:

CC
CC
CC

COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/238,163

CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian Kevin I

CC NAME: BASCLAU, REV III L.
CC REGISTRATION NUMBER: 34 774

CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 543-9

TELEPHONE: (415) 543-5043
TELEFAX: (415) 543-5043

CC INFORMATION FOR SEQ ID NO: 5:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 215 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC
TOPOLOGY: unknown

CC MOLECULE TYPE: protein

CC
FEATURE:

CC NAME/KEY: misc_feature

CC LOCATION: 1..215
 CC OTHER INFORMATION: /standard_name= "Deduced amino acid
 CC OTHER INFORMATION: sequence of PCIP from bean."
 CC SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 2.1%; Score 36; DB 1; Length 215;
 Best Local Similarity 16.7%; Pred. No. 9.86e-08;
 Matches 31; Conservative 69; Mismatches 85; Indels 1; Gaps 1;

Db 10 SVVSRATSCNDKAKKDDNTSSWTTDCNRTGWVCDTDTTYRVNNDGHNKYSANNYG 69
 Cp 955 CAAGATACGAAGCGGAACCGACGTCAGAAACCCGCGATACCAATTAAGTTTAGAGCG 896
 Db 70 GNNVGAATHTYHTNVSADSKTVDTSYNASGTSSSNGGT-DGNRSGADSYGSSKTAMT 128
 Cp 895 CGGTCAGGAGCGTGGCCCTGGTGAGGAGCAATAAAAGGTCTGGATGAAGAGTTGACT 836
 Db 129 SRNRTGTANNVDSRNMGDASVSKNTKKHAKNSADSGKVGSKNNGDRNNRYGTGKSN 188
 Cp 835 CGGCCGAAGATCATGACCGGTAGTAGTCCAGTCTGGTAGCTAGTACTAGGAACCGTGCC 776

189 VSNNGC 194
 : :
 775 ACTTCG 770

RESULT 8
 ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
 AC xxxxxx
 DT

Sequence 5, Application US/08238163
 Sequence 5, Application US/08238163
 Patent No. 5569830
 GENERAL INFORMATION:
 APPLICANT: BENNETT, Alan
 APPLICANT: LABAVITCH, John M.
 APPLICANT: POWELL, Ann
 APPLICANT: STOLTZ, Henrik
 TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
 TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Kourie and Crew
 STREET: Stewart Street Tower, One Market Plaza
 CITY: San Francisco
 STATE: California
 COUNTRY: US
 ZIP: 94105-1493

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/238,163
 FILING DATE: 03-MAY-1994
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Bastian, Kevin L.
 REGISTRATION NUMBER: 34,774
 REFERENCE/DOCKET NUMBER: 2307E-540
 TELEPHONE: (415) 543-9600
 TELEFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 215 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 NAME/KEY: misc_feature

CC LOCATION: 1..215
 CC OTHER INFORMATION: /standard_name= "Deduced amino acid
 CC OTHER INFORMATION: sequence of PCIP from bean."
 CC SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 2.0%; Score 34; DB 1; Length 215;
 Best Local Similarity 13.0%; Pred. No. 1.45e-06;
 Matches 19; Conservative 60; Mismatches 67; Indels 0; Gaps 0;

Db 45 DTDTRYVNDGHNKYSANNYGNNVGAATHTYHTNVSADSKTVDTSYNASGTS 104
 QY 1145 GTTCTTTGGTGGCTGCAGTTCAGTTGGAGCACCACTTGTTCCTAGCTGCCGCTGG 1204
 Db 105 SSGGTGDNRSADSYGSSKTAMTSRRTCKTANNVDSRNMGDASVSGDKNTKKHAKNS 164
 QY 1205 GCACCTTAGGAAGATTGGCCCTTGGCTCGGAGCTTGTGTAAAGACGACGGATGCCGTA 1264

Db 165 ADGKVGSKNNGDRNNRYGTGKSNVS 190
 QY 1265 TAGGAGCTTCGGGTTTGGGACGAGC 1290

RESULT 9
 ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
 AC xxxxxx
 DT

Sequence 94, Application PC/TUS9511934
 Sequence 94, Application PC/TUS9511934
 GENERAL INFORMATION:
 APPLICANT: Cytogen Corporation
 TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
 TITLE OF INVENTION: Peptide Libraries
 NUMBER OF SEQUENCES: 103
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/11934
 FILING DATE: 20-SEP-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 1101-196-228
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELE: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 94:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 74 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

Query Match 1.6%; Score 27; DB 4; Length 74;
 Best Local Similarity 10.1%; Pred. No. 1.16e-02;
 Matches 7; Conservative 20; Mismatches 42; Indels 0; Gaps 0;

Db 3 GAGNN 62
 Cp 686 GAGGTCGGGTCGTAGTTCGAGGCTGTTGCAGCGAGGCTGTGGGCTTGTGGTCCACTT 627

Db 63 BNNBNACGC 71
Cc : : : : :
Cp 626 CCACCACGC 618

RESULT 10
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
Sequence 92, Application PC/TUS9511934
Cc Sequence 92, Application PC/TUS9511934
Cc GENERAL INFORMATION:
Cc APPLICANT: Cytogen Corporation
Cc TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From
Cc TITLE OF INVENTION: Peptide Libraries
Cc NUMBER OF SEQUENCES: 103
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Pennie & Edmonds
Cc STREET: 1155 Avenue of the Americas
Cc CITY: New York
Cc STATE: New York
Cc COUNTRY: USA
Cc ZIP: 10036
Cc COMPUTER READABLE FORM:
Cc MEDIUM TYPE: Floppy disk
Cc COMPUTER: IBM PC compatible
Cc OPERATING SYSTEM: PC-DOS/MS-DOS
Cc SOFTWARE: Patentin Release #1.0, Version #1.30
Cc CURRENT APPLICATION DATA:
Cc APPLICATION NUMBER: PCT/US95/11934
Cc FILING DATE: 20-SEP-1995
Cc CLASSIFICATION:
Cc ATTORNEY/AGENT INFORMATION:
Cc NAME: Mirock, S. Leslie
Cc REGISTRATION NUMBER: 18,872
Cc REFERENCE/DOCKET NUMBER: 1101-196-228
Cc TELECOMMUNICATION INFORMATION:
Cc TELEPHONE: (212) 790-9090
Cc TELEFAX: (212) 869-9741/8864
Cc TELEX: 66141 PENNIE
Cc INFORMATION FOR SEQ ID NO: 92:
Cc SEQUENCE CHARACTERISTICS:
Cc LENGTH: 81 base pairs
Cc TYPE: nucleic acid
Cc STRANDEDNESS: single
Cc TOPOLOGY: linear
Cc MOLECULE TYPE: DNA (genomic)
Cc SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

Query Match 1.6%; Score 28; DB 4; Length 81;
Best Local Similarity 12.5%; Pred.No. 3.36e-03;
Matches 9; Conservative 20; Mismatches 43; Indels 0; Gaps 0;

Db 7 CTCGAGNN 66
Cc : : : : :
Cp 689 CTGGAGTCGGGGTCGTAGTCGAGGCTTTCAGCGGAGTGGTGGGCTTGTGGGTCCA 630
Cc : : : : :
Db 67 BNNBNBNACGC 78
Cc : : : : :
Cp 629 CTTCCACCACGC 618

RESULT 11
ID US-08-273-846-1 STANDARD; DNA; UNC; 290 BP.
AC xxxxxx
DT
Sequence 1, Application US/08273846
Cc Sequence 1, Application US/08273846
Cc Patent No. 5641863
Cc GENERAL INFORMATION:
Cc APPLICANT: SCHREIBER, ALAN D.
Cc APPLICANT: PARK, JONG-GU

TITLE OF INVENTION: METHODS OF STIMULATING PHAGOCYTOSIS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/273,846
FILING DATE: 12-JUL-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 555-29
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE 290 BP; 17 A; 34 C; 8 G; 11 T; 220 OTHER.

Query Match 1.6%; Score 27; DB 2; Length 242;
Best Local Similarity 11.6%; Pred.No. 1.16e-02;
Matches 24; Conservative 41; Mismatches 141; Indels 1; Gaps 1;

Db 2 ALLEDFROMPROCEDURENUCLECFRPROC*FXPCALLEDFROMPROCEDURERFDES 61
Cc : : : : :
Cp 1441 AAGAGATCTCCACCCACACACAAACACATACACATGCGTGTATAGCTTCC 1382
Cc : : : : :
Db 62 CCRFPFROC*FXPCALLEDFROMCFRFORM*FXPCALLEDFROMCFRCONVT*FXPCALL 121
Cc : : : : :
Cp 1381 CCATACCCCAAGTTCTTAGGCGCGCGCGAATTAAGTCAACGCGCTGAACCGCGCA 1322
Cc : : : : :
Db 122 EDFROMON**CALLEDFROMP-ROCEUREGETPATHCCRFPAPPSCRF*EXECALLED 180
Cc : : : : :
Cp 1321 TCCCTCAGCGTCCGAATGTCTGACATTAGCGTCCCAAAACCGAAGCTCTATAC 1262
Cc : : : : :
Db 181 ROMON**CALLEDFROMCMCRFPAPSC 207
Cc : : : : :
Cp 1261 GGCAFCGCGTCTTCTTACACAAAGTCC 1235

RESULT 12
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT
Sequence 94, Application PC/TUS9511934
Cc Sequence 94, Application PC/TUS9511934
Cc GENERAL INFORMATION:
Cc APPLICANT: Cytogen Corporation
Cc TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From
Cc TITLE OF INVENTION: Peptide Libraries
Cc NUMBER OF SEQUENCES: 103
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Pennie & Edmonds
Cc STREET: 1155 Avenue of the Americas
Cc CITY: New York
Cc STATE: New York
Cc COUNTRY: USA
Cc ZIP: 10036

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNTE
CC INFORMATION FOR SEQ ID NO: 94:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

```
Query Match 1.5%; Score 26; DB 4; Length 74;
Best Local Similarity 8.8%; Pred. No. 3.89e-02;
Matches 6; Conservative 20; Mismatches 42; Indels 0; Gaps 0;
```

[illegible]

D5	65	NBNACGCC	72
		:	
QY	217	CCGACGCC	224

RESULT 13
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx

Sequence 100, Application PC/TUS9511934
 CC Sequence 100, Application PC/TUS9511934
 CC GENERAL INFORMATION:
 CC APPLICANT: Cytogen Corporation
 CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
 CC TITLE OF INVENTION: Peptide Libraries
 CC NUMBER OF SEQUENCES: 103

CC
CC
CC
CC
CC
CC

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

REF: 10030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11934
FILING DATE: 20-SEP-1995

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-196-228
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864

```

CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 100:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 74 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
CC      SEQUENCE 74 BP: 6 A; 6 C; 1 G; 1 T; 60 OTHER.
SQ

```

```
Query Match      1.5%; Score 25; DB 4; Length 74;
Best Local Similarity 9.0%; Pred. No. 1.28e-01;
Matches 6; Conservative 20; Mismatches 41; Indels 0; Gaps 0;
```

[illegible]

Db 61 NNVNNAC 67
Cp 1207 TGCCCCAC 1201

RESULT 14
ID PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.
AC xxxxxx

Sequence 99, Application PC/TUS9511934
Sequence 99, Application PC/TUS9511934
GENERAL INFORMATION:
xxxxxxx
AC
DT
DE
CC
CC

CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036

CC ZIP: 10030
CC
CC COMPUTER READABLE FORM:
CC
CC MEDIUM TYPE: Floppy disk
CC
CC COMPUTER: IBM PC compatible
CC
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC
CC CURRENT APPLICATION DATA:
CC
CC APPLICATION NUMBER: PCT/US95/11934
CC
CC FILING DATE: 20-SEP-1995
CC

CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-328
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEEX: 66141 PRNNP

CC	TELEX: 66141 PENNIE	99:
CC	INFORMATION FOR SEQ ID NO:	
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH: 75 base pairs	
CC	TYPE: nucleic acid	
CC	STRANDEDNESS: single	
CC	TOPOLOGY: linear	
CC	MOLECULE TYPE: DNA (genomic)	
CC	SEQUENCE 75 BP: 1: A: 1 C: 7 G: 5	

```
Query Match      1.5%   Score 26; DB 4; Length 75;
Best Local Similarity 9.08; Pred. No. 3.89e-02;
Matches        6; Conservative 20; Mismatches 41; Indels 0; Gaps 0;
```


Search completed: Thu Feb 18 17:54:01 1999
Job time : 240 secs.